Fall 8-11-2020

ESSAYS IN URBAN ECONOMICS

Kalee E. Burns
kaleeb36@gmail.com

ABSTRACT

ESSAYS IN URBAN ECONOMICS

By

KALEE ELISE BURNS

AUGUST 2020

Committee Chair: Dr. Carlianne Patrick

Major Department: Economics

This dissertation examines different topics within the sphere of urban economics. My first chapter, "Social Capital and Entrepreneurship", explores the role of social capital – at the community and individual level – in the pursuit of entrepreneurial opportunities measured through self-employment. I define social capital along two dimensions: trust and network social capital. Using two-sample two-stage least squares (2S2SLS), I construct measures of trust and network social capital for respondents in the restricted 2000 U.S. Long Form Decennial Census and predict the likelihood that a person is observed in a self-employed state using a multinomial regression. I find that individuals with high amounts of network social capital characterized by informal relationships (hereafter referred to as "weak network social capital") are 0.56 percentage points more likely to be observed in self-employment compared to individuals with medium or low levels of weak network social capital. Measuring social capital at the census tract level, only individuals living in communities with high levels of weak network social capital and strong network social capital (characterized by familial or close-friend relationships) have statistically higher probabilities of being observed in self-employment relative to individuals in communities with low and medium levels of network social capital. Stratifying the sample by urbanicity, the relationship between social capital and self-employment is stronger for individuals with high

weak networks who live in the most rural census tracks in the United States. These results imply that increasing social interactions in communities through the promotion of social capital building entities (i.e. clubs and social groups) may be an innovative and low-cost intervention for communities with potentially poor labor market opportunities.

In my second chapter, “Amenities and the College-Educated: A Gentrification Perspective," I examine gentrification across all metropolitan areas in the U.S. between 2007 and 2014 and develop a conceptual framework of gentrification that has causal interpretations. Gentrification is an important topic within public policy, being a subject of debate and interest for economists, city planners, and politicians. Recent evidence of gentrification in several U.S. cities finds a significant correlation between gentrification (via increases in a neighborhood’s share of college graduates) and the location of consumption amenities. This paper develops a conceptual model as the basis for estimating gentrification and amenity establishment location simultaneously for Core-Based Statistical Areas in the U.S. between 2007 and 2014. I use a measure of operationally defined gentrification whereby the underlying measure is share of college graduates in a neighborhood. I find that when controlling for the simultaneous nature between amenities and gentrification, gentrification increases the consumption amenities in a neighborhood by 4%. Contrastingly, there is no clear evidence that gentrification increases the number of neighborhood amenities

Finally, my third chapter, “Revisiting the Burden of the Gas Tax in an Electric Vehicle World," examines changes in the distribution of the gasoline tax burden in the presence of increased electric vehicle adoption. In the last several decades there has been a large growth in the number of electric vehicles on the roads. However, even with this growth, the primary source of infrastructure funding in the U.S. continues to be gasoline taxes. Less demand for gasoline may impact the elasticity of demand for gasoline, therefore potentially threatening a revenue source. Furthermore, the burden of the tax will continue to shift towards consumers of gas-powered vehicles. This chapter re-examines a model of gas-tax incidence using updated consumer data. I then simulate electric vehicle purchases to examine the burden on consumers when the national gasoline tax is increased.

ESSAYS IN URBAN ECONOMICS

BY

KALEE ELISE BURNS

A Dissertation Submitted in Partial Fulfillment of
the Requirements for the Degree
of
Doctor of Philosophy in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2020

ACCEPTANCE

This dissertation was prepared under the direction of Kalee Elise Burns's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair:	Dr. Carlianne Patrick
Committee:	Dr. Julie Hotchkiss Dr. Garth Heutel Dr. Laura Wheeler

Electronic Version Approved

Sally Wallace, Dean
Andrew Young School of Policy Studies
Georgia State University
August, 2020.

DEDICATION

I dedicate this dissertation to my mother, Susan Grice Burns, and my grandmother, Patricia Grice.

ACKNOWELDGEMENTS

I offer my sincerest gratitude to my advisor, Dr. Carlianne Patrick, for her mentorship and guidance. You always made time to give me advice and better my research even when you were busy with so many things. Secondly, I thank Dr. Julie Hotchkiss for taking me under your wing as a research assistant and helping me realize my true passion of working for a government institution. I owe all my current organizational skills to you. Finally thank you to Dr. Garth Heutel and Dr. Laura Wheeler for your time and effort into making me a better economist.

I also would like to express my appreciation for the faculty of the Economics Department at Georgia State University. My research has greatly improved through the comments, lectures, and conversations with Dr. Pierre Nguimkeu, Dr. Dan Kreisman, and Dr. John Gibson. A special thanks is also extended to Bess Blyer and Kristy Hill who were instrumental in making the teaching and job market portion of my graduate experience as smooth as possible. A very special to Dr. Melissa Banzhaf and Dr. Melinda Pitts for all the support and encouragement over the years during lunches at the Fed! Finally, I thank Dr. Richard Hawkins for his mentorship since my time at the University of West Florida

I am grateful to my friends who graduated before me, Cary Cruz Bueno, Pelin Ozturk, Makayla Palmer, and Florian Rundhammer. Your advice and guidance were immensely helpful in preparing and surviving the economic job market. A very special thank you to my three dear friends, Camila Morales, Alicia Plemmons and Mia Goodnature. Your constant and unrelenting support via conversations and detailed proof-reading, and most importantly, friendship, is irreplaceable. My heartfelt appreciation goes to Sei-Youn Kim, Maria Sudibjo, Katherine Key, Franco Beltran, Jordan Jones, Olivia Delmas, and Chandrayee Chatterjee. Thank you to Kaylea, Hillary, Lana, and Gracie for your friendship that spans almost 10 years. Finally, I would like to

thank my family for their constant love and jokes about my constant state of being a student for the better part of a decade. The doctor is in now!

Table of Contents

List of Tables

List of Figures

Chapter I: Social Capital and Entrepreneurship

I. Introduction

Entrepreneurship is a vital part of the U.S. economy. As of 2019, thirty percent of the workforce is either self-employed or working for someone who is self-employed.[1] The self-employed themselves make up ten percent of the workforce in the United States.[2] Individuals' abilities, family background, financial constraints, and strength of ethnic enclaves are all important factors behind an individual's choice to pursue self-employment (Le 1999). This paper explores the potential role of social capital as an additional contributor to the pursuit of entrepreneurial opportunities as measured through self-employment. I find a person's general level of trust and strength of social networks, as well as the community's level of these social capital measures, to be important, in varying degrees, for a person's self-employment decision. Robert Putnam's work, "Bowling Alone" describes how social capital in America has been in a decline since the 1970s. As a potentially important resource in employment decisions, declines in social capital have implications for the United States labor market. The findings in this paper could prove useful to policy makers in communities with a goal of promoting entrepreneurial and other economic activity.

Defined as the value of social connections, formal and informal, that individuals can use to achieve private or collective objectives, social capital is commonly referred to as the societal counterpart to physical or economic capital.[3] This paper will examine social capital's influence on the self-employment decision along two dimensions: structural social capital and cognitive social capital. Harpham et al. (2002, p. 106) characterize the two dimensions: "as what people

---

[1] 10 facts about American workers
[2] Three-in-Ten U.S. Jobs Are Held by the Self-Employed and the Workers They Hire
[3] Social Capital Community Benchmark Survey Description

‘do’ and what people ‘feel’ in terms of social relations.” Cognitive social capital “derives from individuals’ perceptions and mental processes resulting in norms, values and beliefs that promote cooperation” (Sabatini, 2015). Cognitive social capital is usually in the form of social trust or trust in members of one’s own family (Lancee 2010; Kwon, Heflin, and Ruef 2013; Hotchkiss and Rupasingha 2018a). In this paper, I quantify cognitive social capital using various measures of trust. Contrastingly, structural social capital is network-based and quantified by connections in a person’s social network and participation in group activities. Additionally, structural social capital takes two forms: strong and weak. This delineation of ties within a network comes from the work of Mark Granovetter who describes the relative strength of network as either a network connection between two close friends (a strong tie) or a connection between two acquaintances (a weak tie) (Granovetter 1973). Davidsson and Honig (2003) measure strong network social capital, in the context of entrepreneurial activity, as whether an individual has a family member who owns a business. Informal, or weak, ties have often been measured as the number of organizations in which an individual participates or to which an individual belongs (Kwon et al. 2013). For the rest of this paper the two types of structural social capital will be called weak network and strong network social capital.

While previous work has examined the relationship between entrepreneurship and some forms of social capital along both of these dimensions, the generalizability of the studies is limited by sample size and restrictions to a selected portion of the population (see Stam et al., 2014 for a review of the literature). In contrast, I utilize the restricted microdata from the 2000 US Decennial Census (DC) which is a large, nationally representative sample. This survey is a 1 and 10 sample of the US population. Therefore, it contains over 20 million individual observations, with detailed demographic, economics, and most importantly, fine geographic-level information. The geographic nature of this is of importance to how one believes social capital is generated. Because social capital is who a person knows and is surrounded by, geography plays an important role. By utilizing this unique and rich data source I am able to improve upon previous studies' measure of individual and community level social capital by using fine geographic information. More specifics of this will be described in a later section.

Additionally, this paper contributes to the existing literature by making use of a broad range of activities reported in the Social Capital Community Benchmark Survey of 2000 (SCCBS) to construct measures of trust and weak and strong network social capital.[4] Further, I compare the importance of individual-level versus community-level social capital measures for both trust and network measures of social capital in the context of self-employment. Previous studies either use only community social capital or one type of individual social capital (Kwon et al. 2013; Nieto and González-Álvarez 2016); Nieto & González-Álvarez, 2016). The analysis here also improves upon earlier studies by controlling for selection into the labor market and presents a conclusive argument for the empirical relevance of both measures using a large nationally representative data set.

My analysis seeks to answer the following question: How does social capital influence the probability of self-employment relative to other labor market statuses? This question is answered utilizing restricted and confidential data through the combination of both the Census Bureau's Research Data Center and the Roper Center's Social Capital Community Benchmark Survey. Social capital is estimated for respondents in the Decennial Census using two-sample two-stage least squares, allowing me to make use of a large, nationally representative sample for

[4] A more recent 2006 SCCBS survey exists, however it does not contain an indicator for census tract, which is important for constructing an instrument for individual social capital (see https://ropercenter.cornell.edu/2006-social-capital-community-benchmark-survey/).

the analysis. Allowing for three labor market statuses (not in the labor force, paid employment, and self-employment), multinomial logit regression analysis provides results unconditional on the labor force participation decision.

The effect of network social capital at the individual level depends on the strength of the network. Weak network social capital positively predicts self-employment while strong network social capital negatively predicts self-employment. This is consistent with the idea that a weak individual network allows for more new information to flow through the network while a strong network can impede information about possible job opportunities. At the community level, both network social capital measures are positively predictive of a person being observed in self-employment. A person's level of trust, and overall trust exhibited by a person's community, was not an important factor in positively predicting self-employment decisions. These results are consistent across differences in demographics and economic stratifications such as immigrant status, “urbanicity” of community, gender, and educational attainment. Of these various stratifications, the importance of weak network social capital in the prediction of self-employment in some of the most rural communities might offer policy makers a tool to spur job opportunities for individuals facing potentially poor employment outcomes.

The remainder of this paper is as follows: Section II briefly details relevant literature to social capital and entrepreneurship; Section III discusses the sources of data used; Section IV describes the empirical methodology used to connect social capital with entrepreneurship; Section V presents the results; Section VI concludes.

II. Literature Review and Background

II.A. Individual vs. Community Social Capital and Entrepreneurship

Much of the previous literature relating entrepreneurship and social capital focuses on individual or community social capital but not both (for example, see Davidsson and Honig, 2003; Rupasingha et al. 2006; Kwon et al 2013). Nieto and González-Álvarez (2016) examined the joint contribution of community and individual network social capital on entrepreneurship, but for Spain only. One of Glaeser (2007)'s four theories of entrepreneurship suggests that entrepreneurial activity in one sector of a community can create a "culture of entrepreneurship," increasing entrepreneurial activity community-wide. Work by Stam et. al (2014) suggests that weak network social capital (e.g. social capital associated with participation in volunteer groups) is a stronger determinant of entrepreneurship than strong network social capital (e.g. social capital associated with number of close friends and family). Furthermore, results from Nieto and González-Álvarez (2016) suggest that individual social capital likely has more of an impact on entrepreneurship than community level social capital (in Spain). Strong network social capital is typically associated with less influx of information. In Kwon et al. (2013) people who belonged to religious and sports-affiliated types of organization are less likely to be connected to any other organization. Due to this restriction on dissemination of new information, members with strong network social capital may be less aware of entrepreneurial opportunities and are therefore less likely to pursue self-employment. Or, alternatively, due to the close nature of their relationships with members in their (sports and religious) network, they may be better information about paid employment opportunities. Therefore, it is possible that a higher level of this type of social capital decreases the probability that an individual becomes an entrepreneur. However, if someone is interested in entrepreneurial pursuits and belongs to a close-knit community, then this may be advantageous to their ability to exploit business connections.

The analysis here contributes to this body of knowledge by combining what has only been investigated in pieces and, as far as I know, is the first to do so with a nationally representative sample for the U.S. Specifically, both trust and network measures of social capital are considered, and both types of social capital are measured at both the individual and community level. A "community" in this paper is measured at the Census tract level. This is a more geographically concentrated measure of community than in previous literatures (i.e., Kwon et al. 2013).

II.B. Nativity and Entrepreneurship

This paper also explores how social capital might differentially impact entrepreneurial choice decisions between immigrants and natives. In the U.S. labor force, immigrants make up a disproportional share of entrepreneurs. According to Kerr and Kerr (2016), approximately 24% of entrepreneurs are immigrants but constitute only 19% of the U.S. workforce. Therefore, entrepreneurship appears to be an important form of employment among immigrants.

In his investigation of the effect of social capital on employment status among four groups of immigrants to the Netherlands, Lancee (2010) finds that weak network social capital is a better predictor of self-employment than strong network social capital. Borjas (1986) attributes the higher rates of entrepreneurship among immigrants to an "enclave" effect which increases self-employment opportunities due to the concentration of residents in a community along a shared cultural identity. This shared cultural identity can be a contributing factor to an immigrants' weak network. My research also supplements work by Kerr and Kerr (2016), who examine differences in entrepreneurship between native and immigrants. In that work, the authors find that immigrant entrepreneurs have better three- and six-year employment growth outcomes than natives. If different types of social capital (trust vs. weak/strong network) and from different sources (individual vs. community) impact self-employment decisions differently between natives and immigrants, communities may be able to tailor their investments in social capital in ways that will be most effective depending on the demographics of the community.

II.C. Rural vs. Urban Differences in Entrepreneurship

The potential importance of community level social capital in the determination of entrepreneurship suggests that other characteristics of that community might influence the interaction between social capital and entrepreneurship. In an urban setting, it is plausible to think that one has more weak network social capital due in part to casual contact with more people. Contrastingly, in a rural setting one's interactions are likely to mostly involve contact with family and friends.

Lannoo et al. (2012) lists several reasons why the relationship between social capital and entrepreneurship might differ between urban and rural communities. For example, these authors suggest urban residents have shallow or superficial social connections which would be expected to result in weaker network social capital compared to rural residents. While the authors do not specifically speak to entrepreneurial decisions from differences in social capital along urban and rural lines, the fact that urban and rural residents might have different levels of social capital (of different types) speaks to the importance of investigating its effects in different environments. Stam et al. (2014) find that weak network social capital is more predictive of entrepreneurship than strong network social capital; their work suggests that one would find more entrepreneurial activity in an urban community.

Additionally, in urban areas there are a plethora of institutions that a potential entrepreneur can take advantage of when pursuing entrepreneurial endeavors, such as banks and business networking events. However, rural areas are often lacking in these types of institutions. Social capital may act as a substitute for these types of institutions in small communities or rural areas. This is the theory that Bauernschuster et al. (2010) posit in their article. Using information on individuals' club memberships, the authors find that there is a larger effect of weak network social capital on the probability of individuals becoming an entrepreneur in communities with less than 5,000 people. In addition to adding to the evidence of how the relationship between weak and strong network social capital and entrepreneurial activity differs across urban and rural settings, the analysis here will also identify differences in the role community versus individual social capital plays in a rural versus urban setting.

In the context of employment, self-employment is particularly important for economic growth in rural areas. Given that rural and non-metropolitan areas in the United States face more instances of paid employment constraints and higher rates of underemployment (Henderson 2002; Findeis and Jensen 1998), social capital can act a particularly valuable resource in the pursuit of self-employment opportunities.

III. Data

This paper utilizes two nationally representative restricted data sources, the 2000 Social Capital Community Benchmark Survey (SCCBS) and the long form 2000 U.S. Decennial Census (DC). The SCCBS was administered through the Roper Foundation in 2000 and approximately 25,000 people responded to the survey. One aspect of the survey that makes it a good fit for matching with other restricted data sources such as the DC is the record of census tracts, which are an often-used measure of community or neighborhood. The SCCBS contains an array of questions related to an individual's social capital, as well as demographic and household economic information. It does not however include labor market status information outside of whether a person is in the labor force or not.

The long form DC was administered to approximately 10% of US households. Unlike the short form, the long form contains information outside of basic demographics, such as detailed labor market information. The DC contains millions of observations. While the DC is much larger than the SCCBS it is still possible to use the SCCBS to predict social capital through a process described in the next section because both contain similar information. Summary statistics for the DC will be discussed below.

IV. Methodology

Social capital and entrepreneurship are not both observed in a single survey or data set. Therefore, everyone in the DC must be assigned an estimated measure of their social capital of different types. The next three subsections describe the methodology employed to calculate both individual and community levels of trust social capital, and weak and strong network social capital. The last two subsections describe the methodology used to investigate social capital's influence on entrepreneurial outcomes.

Two-Sample Two-Stage Least Squares (2S2SLS) is commonly used to overcome the problem of missing data in the primary data set. This type of instrumental procedure was first developed by Klevmarken (1982) as an estimation strategy for variables that are not contained within a single sample (also see Ridder and Moffitt 2007). One of the most well-known applications of two-sample instrumental variable estimation comes from Angrist and Krueger (1992). In their analysis, Angrist and Krueger estimate the effect of age at school entry on educational outcomes. Since age at school entry and educational attainment are not contained within a single sample, the authors employ two-sample instrumental variables. This predicament is similar to the one faced here -- the DC does not contain information on social capital but does contain the outcomes of interest, namely the labor market status of an individual. The Social Capital Community Benchmark Survey (SCCBS), described above, is used to obtain predictors of social capital at the individual level that will be applied to observations in the DC. This multi-step procedure is described below.[5]

IV.A. Estimating Social Capital for Observations in the DC and CPS

The basic strategy is to estimate a social capital determining equation using the SCCBS, then use those parameters to predict multiple dimensions of social capital for observations in the DC. The next three sub-subsections describe the three steps needed to estimate social capital.

IV.A.i. Reweighting the SCCBS

The first step in applying Two-Sample Two-Stage least squares (2S2SLS) to get parameters with which to predict social capital in the DC is to make sure both the predicting data set (SCCBS) and the data set for which predictions are made (DC) have a common set of variables and to construct weights for use in the estimation that make the two data sets look similar (at least at the means). Inverse probability weighting (see DiNardo et al. 1996) is used to make the SCCBS more similar to the DC. The SCCBS is appended to the DC and the following equation is estimated using logistic regressions:

$$P(observation\ i \in SCCBS\ |X) = \Lambda(X'b). \qquad (1)$$

The parameter estimates from this regression are then used to construct the inverse probability ratio, $\frac{\Lambda(X'\hat{b})}{1-\Lambda(X'\hat{b})}$, for each observation in the SCCBS. This is the re-weighting function used to modify the individual weights provided in the SCCBS.[6]

[5] The 2S2SLS methodology described in this analysis follows the methodology implemented by Hotchkiss and Rupasingha (2018). The authors in this paper measured the determinants of social capital in a similar fashion but estimate different types of social capital than are estimated in this paper.
[6] Parameters from the re-weighting exercise have not been disclosed but don't show anything out of the ordinary.

IV.A.ii. Identifying a Person's Unobserved Social Capital through Factor Analysis.

Since social capital is not a characteristic that is observed, the next step is to create measures of social capital for individuals in the SCCBS. This is done using factor analysis designed to capture a common social capital characteristic (e.g., weak network, strong network, or trust) exhibited by an individual participating in certain activities or exhibiting particular attitudes. For example, questions related to generalized trust, trust of people in their neighborhood, and trust amongst working colleagues, among others, are used to construct a person's level of trust social capital. Questions related to membership status in various organizations such as in art groups, neighborhood associations, as well as information on whether an individual volunteered in activities that are unrelated to membership in any group, are used to construct a person's level of weak network social capital. A measure of strong network social capital is constructed from information about connections and interactions with friends and families. Table I details each of the components that are combined, through factor analysis, to construct a single measure for each social capital characteristic.

Factor analysis elicits the common factor from responses to multiple questions in the SCCBS about activities related to a specific measure of social capital in order to uncover a person's latent degree of social capital. From this analysis, I obtain a single factor values for each person for each type of social capital. For each measure of social capital, the factor is essentially a linear combination of the original variables (responses to survey questions) combined to reflect the person's latent social capital level.

IV.A.iii. Estimating the Determinants of Social Capital.

Once measures for each type of social capital for each person in the SCCBS are obtained, the way in which observable characteristics are related to the measured level of social capital are obtained. These social capital "factor" values are not easily interpretable, but they are ordinal by construct. So, the distribution of their values is split into three levels--high, medium, and low--which are then used to estimate an ordered logit to determine the relationship between a person's observable characteristics and their level of each type of social capital. For all types of social capital considered, the process of estimating social capital is the same.

The probability that individual i, living in census tract c, has social capital level k (low, medium, or high) of type j (j = trust, weak network, strong network) is formally expressed as:

$$P\left(SK_{ic}^{j} = k\right) = Pr(\mu_{k-1} < \alpha_0 + \gamma_1 X_{ic} + \gamma_2 Z_c + \mu_{ic}), \tag{2}$$

Where SK_{ic}^{j} is the social capital of type j of person i who lives in census tract c. X_{ic} are individual characteristics such as race, ethnicity, marital status as well as regional dummies. Z_c are regressors that are unique to the equation. These include the distance-weighted average of nearby census tract characteristics, such as the labor participation rate, the percentage of homeowners, the percent of households with children in the home, etc. The full list of Census tract controls is in Appendix Table AI. These census tract variables are constructed using the 2000 DC. In order to address the potential endogeneity of the characteristics in which an individual lives and their level of social capital, Z_c actually includes the average of these characteristics in surrounding census tracts inversely weighted by distance from an individual's own census tract. The parameters from the ordered logit are then used to estimate where each respondent in the DC lies in each social capital measure's distribution (whether the respondent is expected to have a low, medium, or high level of trust, for example). For the full set of parameters in both X_{ic} and Z_c see Appendix Table AI.

IV.B. Probability of Entrepreneurship

The first analysis assesses the role of individual and community social capital, along with other individual and community characteristics, in observing self-employment, paid employment, or not in the labor force:

$$P(WorkerClass_{ic} = m) = f(X_{ic}, SK_{ic} = high, \overline{SK}_c = high, \bar{X}_c, \varepsilon_{ic})\ , \quad (3)$$

where *WorkerClass* is the labor market status of person *i* in census tract *c* (m = 1 is not in labor force, 2 is in labor force in paid employment, and 3 is in the labor force and self-employed); X_{ic} is a $l\ x\ 1$ vector of individual characteristics, such as race, education, and measures of household income; $SK_{ic} = high$ is equal to one if the person has a level of social capital that is in the top third of the distribution for that social capital measure; $\overline{SK}_c = high$ is equal to one if the community's average level of social capital is in the top third of the distribution of that social capital measure across communities*;* and $\bar{X}_c$ are average census tract characteristics, such as the unemployment rate. Community level social capital is a distance-weighted average of individual social capital measures in *surrounding* census tracts, reducing any potential endogeneity with one's individual social capital. Table II list summary statistics for variables used to estimate equation (3). To avoid likely collinearity between the types of social capital, equation (3) is estimated separately for each type of social capital. For instance, there are elements of inherent trust in a person's strong network. Additionally, a person may be trusting and, therefore, more likely to form a large weak network. I focus on levels of high social capital in order to ease interpretation of its effects compared to low and medium social capital.

It is assumed that, conditional on observed community and individual characteristics, a person chooses the labor market status that maximizes their utility. An individual choosing self-employment will do so because it dominates the other options -- paid employment and being out of the labor force (see Patrick et al. 2016). By estimating equation (3) as a multinomial logit model, the correlation between the utility from the three options is accounted for. This approach also means that regressors unique to a selection equation in a standard Heckman selection model are not required; it is often difficult to justify the exogeneity of such regressors. Equation (3) is also estimated separately for natives and immigrant individuals, individuals in rural, suburban, and urban tracts, gender, and education.[7]

The analysis here examines similar types of social capital using similar measurement strategies (types of questions considered) as Kwon et al. (2013), however it differs in three ways. First, I perform factor analysis on groups of questions in order to derive a measure of the relevant types of social capital. This allows me to quantify a person's latent social capital. Second, the analysis incorporates both individual and community social capital measures. Finally, whereas Kwon et al. (2013) assigns community average responses (at the PUMA-level) from the SCCBS to the decennial census individual respondents, I estimate social capital measures for each individual separately, based on their individual characteristics and then aggregate to the census tract (community) level. This improves on a community measure of social capital because community social capital is directly related to all members of that community (every person in the Decennial Census) at a more targeted geographic area (a census tract) as opposed to the average responses of a small amount of people in a very large geographic area (SCCBS respondents in their corresponding PUMA).[8]

---

[7] Rural, suburban, and urban census tract status are determined by a census tract's RUCA code from the ERS. See here for more information: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/

[8] The prediction of social capital is off of a small sample, but that sample is reweighted to better match the distribution of the larger sample, therefore still improving on the estimates of social capital for individuals in the Decennial Census.

V. Results and Discussion

V.A. First Stage: Social Capital Estimation

The first-stage ordered logit results corresponding to equation (2) are found in Appendix Table AI. These results tell us how observed individual and census tract characteristics are related to a person's measure of each type of social capital. I identify variation in social capital utilizing exogenous census tract characteristics such as distance-weighted average of nearby Census Tracts' share of jobs in social occupations, the unemployment rate, and population density. Each column of coefficients in Appendix Table AI represents a separate regression. For trust, the statistically significant identifying census tract characteristics are the unemployment rate, median age, and share of families with children. The share of workers in "social" industries and the share of people who lived in the community at least 5 years identifies weak network social capital. Finally, strong network social capital is identified by the unemployment rate, the female labor force participation rate and the share of college graduates. For each social capital equation, the pseudo R^2 is within the range of goodness of fit reported by Hotchkiss and Rupasingha (2018) from their identical ordered logit. Others using an ordinary least squares methodology in the first-stage report goodness of fit measures between 0.01 and 0.604 (Currie and Yelowitz 2000; Dee and Evans 2003; Nicoletti and Ermisch 2008; Cavaglia 2015).

At the individual level, one of the most striking statistically significant determinants of trust is the role of race. Minorities are statistically less trusting than White, non-Hispanics (NH). This is consistent with differences in trust between White, NH and minorities as detailed in Price (2012). Additionally, individual educational attainment (not accounting for the interaction terms) suggests increases the amount of social capital a person has? while recent migration to a location decreases the amount of social capital one has, consistent with Rupasingha et al. (2006).

The coefficients in Appendix Table A-I are what is applied to the DC in order to construct an estimate of unobserved social capital.

V.B. Second Stage: Social Capital and Self-Employment in the Full Sample

Tables III through VIII present the multinomial logit average marginal effects (AME) from estimating equation (3) separately for each social capital measure, for the full sample and then for different sub-samples of the population. I present the AME as opposed to the regression coefficients due to multinomial logit regression coefficients being sensitive to the chosen base category. For completeness, average marginal effects for all labor market outcomes are reported. I first consider the results for the primary sample, which includes all civilian persons between the ages of 18-64. This sample has approximately 20,540,000 persons. Individuals are split among the three worker classes: Not Working (~3,852,000 million), Paid-Employed (~15,040,000 million), and Self-Employed (1,668,000).

V.B.i. Trust and Self-Employment

Focusing first on the role of trust, an individual with a high (relative to medium or low) level of trust social capital has a lower probability of being observed in self-employment by 0.31 percentage points as seen in Table III. At first glance, this might seem counterintuitive. However, individuals with high levels of trust may believe that society will take care of them in a non-working state (hence, more likely to be non-working). The marginal effect for self-employed is significantly smaller than that of paid-employment, suggesting that while individuals with high trust are less likely to be in either paid or self-employment, those who *are* working are more likely to be self-employed than in paid employment. These results illustrate why accounting for selection into the labor market is important. If the analysis excluded non-workers, we would conclude that promoting high levels of individual trust social capital would increase entrepreneurial activity. However, doing so would actually increase not working at the cost of both paid and self-employment.

Community trust does not appear to be a strong predictor of self-employment. This result is in direct contrast to Kwon et al. (2013), who finds a significant and positive association for trust in observing a person in a state of self-employment. One possible reason for this discrepancy is that here a community is a census tract, whereas Kwon et al. measure the community at the PUMA level, a much larger, and less precise, measure of community. Additionally, Kwon et al. do not include individual levels of social capital and do not account for the possibility of not working as a labor market choice.

V.B.ii. Weak Network and Self-Employment

Table IV presents the average marginal effects of weak network social capital on employment status. In contrast to trust, those with high levels of weak network social capital are more likely to be in both paid and self-employment, relative to those with medium or low levels of weak network social capital. A person with a large weak network is likely to be involved in many social groups or spend time volunteering. A large weak network provides a mechanism for information to more freely flow, increasing knowledge of potential employment/business opportunities. Additionally, a large weak network could also act as a consumer base for an entrepreneur. Since the marginal effect of high individual weak network social capital for paid employment (0.0087) is higher than the marginal effect for self-employment (0.0065), this suggests that weak networks may prove more marginally more valuable to finding paid employment opportunities than self-employment opportunities.

The story is different for individuals who are in *communities* with high levels of weak network social capital. An individual who lives in a community where it is common to be very involved in social activities and volunteering is 0.53 percentage points more likely to be self-employed than someone in a community with medium or low levels of weak network social capital. However, that individual is no more or less likely to be in paid employment or not working. A high level of weak network social capital within the community appears to provide an individual with resources important for promoting entrepreneurial activity, such as a more connected community acting as a potential customer base. This result is consistent with the results from Kwon et al. (2013), who also find a positive effect of weak network social capital on self-employment.

V.B.iii. Second Stage: Strong Network and Self-Employment

A person's strong network is quantified by the number of close friends and family they have as well as the nature in which that person interacts with these close relationships. The results in Table V suggest that individuals with a high (vs. medium or low) level of strong network social capital are more likely to be not working and less likely to be both self-employed and in paid employment, much like the effect of high levels of trust social capital. It would make sense that high levels of trust and high levels of strong network social capital might have similar effects, as strong family and friend ties might tend to make someone more trustful. These results suggest that a large strong network provides more support for not working than either employment outcome.

Relative to paid employment, individual high levels of strong network social capital is associated with a higher probability of self-employment -- high levels of strong network social capital reduces the probability of self-employment by about one percentage point, whereas it reduces the probability of paid employment by nearly eight percentage points. This suggests that *if working* a large strong network offers more support/resources for pursuing self-employment rather than paid employment.

Again, juxtaposing these results for strong network with those for trust social capital, we see an interesting difference. Like higher levels of community trust social capital, higher levels of community strong network social capital reduce the probability of not working. However, whereas that reduction in the not working probability as community trust increases is absorbed by a rise in the probability of paid employment, higher community strong network social capital increase self-employment. This suggests that a strong networked community can offer a more stable environment for a person wanting to pursue entrepreneurial activities such as self-employment. In other words, community trust is more important in seeking paid employment, whereas a strong network within the community is more important in promoting self-employment.

V.B.iv. Other Determinants of Labor Market Status

In addition to a person's individual or community level social capital, labor market status is influenced by a number of other factors. This section briefly discusses how these factors influence a person's worker class as seen in Tables III-V. Consistent with Kerr and Kerr (2016), immigrants are more likely to be self-employed than natives. Females are much less likely than males to be self-employed, a finding consistent with Patrick et al. (2016). While there is often a belief that the older one is the less likely they will have a successful business, current literature finds that success of an entrepreneur is positively correlated with age (Azoulay et al. 2018). Here I also find increases in age increase the likelihood of being observed in self-employment. White, non-Hispanics are more likely to be self-employed than minorities. Homeownership and a college education also increase the likelihood an individual is self-employed.

In addition to community social capital measures, other community measures such as the unemployment rate and population density are found to matter in the likelihood an individual is observed as self-employed. Higher unemployment rates decrease the likelihood an individual is self-employed while population density increases that likelihood. While higher unemployment rates indicate an individual is less likely to enter self-employment, the marginal effect is significantly smaller in magnitude than the corresponding ME for paid-employment. Given that this study is for 2000 and the US was at the tail end of an expansion, it is plausible that under more adverse macroeconomic effects such as a recession, that increases in the unemployment rate would increase the likelihood of self-employment relative to paid-employment.

V.C. Second Stage: Social Capital and Self-Employment in Population Subsamples

This section stratifies the full sample across various demographic and geographic dimensions: immigrant status, urbanicity of community, gender, and educational attainment. The empirical strategy is unchanged. Furthermore, for brevity, I only report average marginal effect of the social capital variables. For each of the subsample analyses I estimate a separate regression for each social capital type but only focus on the most striking differences between the various subsamples for discussion.

V.C.i. Immigrants vs. Natives

Immigrants are defined as individuals who are neither a U.S. citizen nor a naturalized U.S. citizen. Table VI presents the marginal effects of each type of social capital on labor market status probabilities for the immigrants and natives separately. The marginal effects for natives for all social capital types, at both the individual and community levels, are consistently similar to the marginal effects for the full sample seen in Tables III-V. Immigrants with high levels of weak network social capital have a statistically larger likelihood of self-employment, relative to their counterparts with medium or low weak levels, than natives (1.3 percentage points vs 0.42 percentage points). In fact, for all employment states with respect to high individual weak network social capital, the marginal effect is larger for immigrants than natives. Immigrants' weak network may be acting as a substitute to institutions that may be unavailable to themselves. In his study of immigrant populations in the Netherlands, Lancee (2010) also finds weak network social capital positively associated with employment outcomes. Additionally, the aforementioned study also finds no connection between trust and employment outcomes.

The lack of statistical significance of all community measures of social capital for immigrants is puzzling. Perhaps the census tract, which is the level at which I measure community, is not the appropriate geography to capture an immigrants "community." Immigrant enclaves are often in very large cities and therefore a census tract in these areas are quite large which may be why community network social capital measures are statistically insignificant.[9]

V.C.ii. Urban vs. Rural

Results in Table VII come from stratifying the sample on the urban and rural status of the community in which a person lives.[10] These results show persons in urban communities generally follow the standard observational pattern as the full sample. The most striking difference between urban and rural residents is related to weak network social capital. Higher levels of both individual and community weak network social capital are much more strongly associated with a higher probability of self-employment among rural residents than among urban residents, indicating that a high individual and community weak network is much more valuable

[9] U.S. Immigrant Population by Metropolitan Area

[10] Urban/Rural status is determined by its Rural Urban Classification Area codes from the Economic Research Service. Urban census tract is one that has a RUCA code from between 1 and 2. A rural census tract has a RUCA code between 9 and 10. Generally, urban census tracts are also tracts in MSAs while rural census tracts are generally outside of MSAs.

to a person in a rural environment than an urban environment. Therefore, while a person in an urban setting is 0.6 percentage points more likely to be self-employed if they have a large weak network, a person in the rural community is 1.2 percentage points more likely to become self-employed. While, in magnitude much smaller, this effect is consistent with Bauernschuster et al. (2010) mark-up effect of weak network social capital. Differences in urbanicity also show that rural communities with large strong networks increase the likelihood of self-employment while those same types of communities in urban areas do not. More so, the marginal effects across all labor market statuses is much larger (as large as 4 times the size) in rural communities with respect to a community with high strong network social capital. Again, strong connections in rural communities, with potentially fewer formal institutions, appears to be an important tool in entrepreneurial pursuits when other options, such as paid employment, may be harder to come by. In other words, social capital is potentially more salient and valuable in the most rural communities.

V.C.iii. Women vs. Men

Table VIII stratifies the main sample by gender. The most striking difference between men and women is the opposite importance of individual vs. community levels of weak network social capital. Whereas high levels of individual weak network social capital are more important than community level weak network social capital in increasing the probability of self-employment among men, the opposite is true for women. The relatively larger importance of community weak network social capital for women is consistent with Patrick et al. (2016) who find that community gender attitudes and the local business climate, generally, play important roles in women's decision to become self-employed. These results are also consistent with Allen

(2000), who finds women “receive less influential social support (akin to this works measure of network social capital) for entrepreneurial activity.”

The importance of community for women's employment status decisions also shows up in the importance, relative to men, of community level trust and community level strong network social capital. Whereas high levels of both community trust and strong network social capital are important for women choosing between paid employment a non-working, they are unrelated to that choice for men.

V.C.iv. College vs. non-College Graduates

Finally, Table IX details the results by individuals with a college degree and individuals without a college degree. The most notable difference between non-college and college graduates is in the role that weak network social capital plays in determining self-employment. High levels of both individual and community weak network social capital are more important among college graduates than among non-college graduates. The college environment provides individuals opportunities to build long-term networks through various groups and volunteering organizations.

For non-college educated individuals living in a community with a large amount of strong network social capital, the probability of self-employment is larger than for college educated individuals in the same type of community. Less educated individuals face more migration constraints and therefore less likely to move (Molloy et al., 2011). As a result, the value of strong bonded community to non-college adults appears to be a stronger predictor of self-employment as shown by a larger marginal effect for this group.

VI. Conclusion

I examine the role of both individual and community level (census tract) social capital in the determination of self-employment as well as other labor market outcomes. The methodology used here accounts for selection into the labor force, making the results more generalizable. The restricted Social Capital Community Benchmark Survey is used to predict three types of unobserved social capital for respondents in the restricted US 2000 Decennial Census using a two-sample two stage least squares framework: trust, weak network, and strong network. By utilizing the richness of the Decennial Census, the analyses in this paper covers over 20 million people and is nationally representative. This improves upon prior literature that primarily focused on targeted populations in order to study the relationship between social capital and entrepreneurship (see Stam et al., 2014). Furthermore, I am able to measure social capital at a finer geographic level than has been previously measured due to the nature of my data.

I find most measures of social capital are statistically significant in determining a person's self-employment status (vs. paid-employment and not working). Individuals with high levels of trust are found to be less likely to enter self-employment, relative to not-working, but relatively more likely to be self-employed relative to paid employment. However, individuals living in *communities* with high levels of social trust are not found to be statistically more likely to enter self-employment.

Both individual and community measures of weak networks are found to statistically increase the likelihood of self-employment. These results are consistent with the way weak networks generally operate: the larger one's weak network (consisting of casual contacts), the more information is available to exploit potential business opportunities. This is particularly true for women, college graduates, and rural residents. Meanwhile, a person with a large strong network is less likely to be self-employed (and less likely to be in paid employment), and this holds across all subsamples. A strong network consisting of family and friends may make the value of not working higher than being in either paid or self-employment. Contrastingly, a community that is very close-knit can provide moral support to an individual pursuing the risky endeavor of self-employment, and, in fact, the results indicate that high levels of community strong networks increases the likelihood of self-employment. This result holds in the subsamples for natives, rural residents, females, and both college and non-college graduates.

The most consistent result across various economic and demographic stratifications is importance of weak network social capital. When thinking about potential policy implications of this work, one can't help but see a potential connection between formation of social capital in rural communities and better employment opportunities for people in these communities. Self-employment in non-metro/rural areas is particularly important for economic growth (Stephens and Partridge 2011). Policy makers considering potential place-based policy in these regions can use the importance of weak networks in their policy designs in order to spur self-employment. The importance of high levels of weak social capital (at both the individual and community level) in increasing self-employment is nearly twice as high in rural communities versus urban communities. The importance of weak network social capital (both at the individual and census tract level) in self-employment decisions in rural areas might offer a promising option for policymaking in terms of how to potentially address the poor paid-employment prospects in rural America.[11] Increasing social interactions in communities through the promotion of social capital building entities like clubs and social groups etc. may be an innovative and low-cost intervention for communities with few paid job opportunities.

[11] Rural Employment and Unemployment

Future work will expand this analysis, using the 2000 Current Population Survey March Supplement, to examine how social capital influences the transition from paid employment to self-employment. This extension is motivated from work by Hotchkiss and Rupasingha (2018b) who find that social capital is influential in the decision behind occupational choice. This analysis will explore the relative importance of previous work experience and social capital in the determination of entrepreneurial activity. In observing an individual's transition from paid employment to self-employment, I will be able to explore how social capital influences people entering entrepreneurship after previously being in a paid employment position. This supplements the current analysis by examining how social capital predicts who *becomes* an entrepreneur.

Tables

Table I: Factor Analysis of Social Capital Questions

Trust	**Weak Network**	**Strong Network**
"Can trust others"	"Participate in charity or social welfare organization"	"Number of close friends"
"Trust neighbors"	"Participate in professional, trade, farm, or business organization"	"Number of people you can confide in"
"Trust co-workers"	"Participate in political group"	"Number of times you've had friends over to your home"
"Trust shop clerks"	"Participate in literary, art, or musical group"	"Number of times you have visited with relatives"
"Trust local police"	"Participate in hobby, investment, or garden club"	"Number of times you have hung with friends in a public place"
"Trust local news"	"Participate in neighborhood association"	
Composite racial trust	"Frequency of times volunteered"	

*These are questions from the Social Capital Community Benchmark Survey

Table II: Chapter I Summary Statistics

Variable	Not Working	Paid Employment	Self-Employed
Social Capital			
High Weak Network- Individual	0.17	0.2652	0.342
	[0.3757]	[0.4414]	[0.4744]
High Weak Network- Community	0.04047	0.04791	0.05009
	[0.1971]	[0.2136]	[0.2181]
High Strong Network- Individual	0.3582	0.3212	0.2053
	[0.4795]	[0.467]	[0.4039]
High Strong Network- Community	0.1159	0.1348	0.1559
	[0.3201]	[0.3416]	[0.3628]
High Trust Individual	0.235	0.3071	0.4218
	[0.424]	[0.4613]	[0.4938]
High Trust Community	0.1902	0.2212	0.2364
	[0.3925]	[0.4151]	[0.4249]
Individual Controls			
Income greater or equal to $30,000	0.6888	0.7864	0.8037
	[0.463]	[0.4098]	[0.3972]
Married	0.5614	0.6049	0.7518
	[0.4962]	[0.4889]	[0.432]
Age	38.74	38.82	44.16
	[13.91]	[11.71]	[10.4]
Age Squared	1695	1644	2058
	[1135]	[933.4]	[909.3]
Immigrant	0.1578	0.1121	0.1159
	[0.3645]	[0.3155]	[0.3201]
Live in area less than 5 Years	0.3364	0.3227	0.2474
	[0.4725]	[0.4675]	[0.4315]
Female	0.5897	0.4925	0.3518
	[0.4919]	[0.4999]	[0.4775]
Black, NH	0.1288	0.09135	0.03989
	[0.335]	[0.2881]	[0.1957]
White, NH	0.6558	0.7589	0.842
	[0.4751]	[0.4277]	[0.3648]
Hispanic	0.1495	0.09629	0.06538
	[0.3566]	[0.295]	[0.2472]
Asian, Other, Non-Hispanic	0.06594	0.05346	0.05277
	[0.2482]	[0.2249]	[0.2236]

Variable	Not Working	Paid Employment	Self-Employed
Less than High School	0.252	0.1301	0.1204
	[0.4341]	[0.3364]	[0.3254]
High School	0.3161	0.2946	0.2741
	[0.4649]	[0.4558]	[0.446]
Some College	0.2825	0.3175	0.2995
	[0.4502]	[0.4655]	[0.458]
College Grad	0.1495	0.2578	0.3061
	[0.3565]	[0.4374]	[0.4609]
Owns Home	0.6931	0.7467	0.8222
	[0.4612]	[0.4349]	[0.3824]
Census Tract Controls			
Unemployment Rate	0.02114	0.01937	0.01846
	[0.01284]	[0.0113]	[0.01056]
Median Household Income	$49,960	$54,210	$59,030
	[21590]	[21600]	[26620]
Population Density	2.092	1.715	1.477
	[5.344]	[4.567]	[4.552]
Number of Observations	3,852,000	15,040,000	1,668,000

*Note means are unweighted. Individuals are ages 16-64 and not classified as working in the military. Sample sizes are approximate in keeping with disclosure guidelines.

Table III: Multinomial Logit- Average Marginal Effects; Trust

	Trust		
Variables	Not Working	Paid Employment	Self-Employed
Social Capital			
High Individual Social Capital	0.0141***	-0.0110***	-0.0031***
	[0.0032]	[0.0031]	[0.0007]
High Community Social Capital	-0.0163***	0.0137**	0.0026
	[0.0047]	[0.0059]	[0.0028]
Individual Characteristics			
Household total income GE $30,000	-0.0433***	0.0484***	-0.0051***
	[0.0016]	[0.0020]	[0.0006]
Married	0.0106***	-0.0289***	0.0183***
	[0.0012]	[0.0016]	[0.0012]
Age	-0.0242***	0.0125***	0.0117***
	[0.0007]	[0.0008]	[0.0002]
Age squared	0.0003***	-0.0002***	-0.0001***
	[8.84e-06]	[9.49e-06]	[2.50e-06]
Immigrant	0.0128**	-0.0272***	0.0144***
	[0.0056]	[0.0053]	[0.0010]
Lived in area 5 yrs. or less	0.0232***	-0.0089***	-0.0143***
	[0.0017]	[0.0015]	[0.0005]
Female	0.0760***	-0.0337***	-0.0423***
	[0.0025]	[0.0025]	[0.0013]
Black, non-Hispanic	0.0135***	0.0156**	-0.0291***
	[0.0046]	[0.0065]	[0.0023]
White, non-Hispanic	-0.0393***	0.0246***	0.0147***
	[0.0040]	[0.0065]	[0.0029]
Hispanic	-0.0036	0.0225***	-0.0189***
	[0.0039]	[0.0051]	[0.0026]
High school education	-0.0530***	0.0530***	0.0000574
	[0.0015]	[0.0018]	[0.0008]
Some college education	-0.0707***	0.0645***	0.0062***
	[0.0017]	[0.0020]	[0.0009]
College graduate	-0.0956***	0.0777***	0.0180***
	[0.0020]	[0.0026]	[0.0012]
Own home	-0.0026***	-0.0088***	0.0114***

	Trust		
Variables	Not Working	Paid Employment	Self-Employed
	[0.0009]	[0.0016]	[0.0010]
Census Tract Controls			
Unemployment Rate	0.5355***	-0.4876***	-0.0479**
	[0.0264]	[0.0325]	[0.0188]
Median Household Income	-1.78e-07***	-3.77e-07***	5.55e-07***
	[3.56e-08]	[4.03e-08]	[1.88e-08]
Population Density	0.0008***	-0.0011***	0.0004***
	[0.0002]	[0.0002]	[0.0001]

Note: Regressions include occupation and MSA fixed effects.
Robust standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table IV: Multinomial Logit-Average Marginal Effects; Weak Network

	Weak Network		
Variables	Not Working	Paid Employment	Self-Employed
Social Capital			
High Individual Social Capital	-0.0152***	0.0087***	0.0065***
	[0.0020]	[0.0020]	[0.0011]
High Community Social Capital	-0.0038	-0.0015	0.0053***
	[0.0043]	[0.0053]	[0.0014]
Individual Characteristics			
Household total income GE $30,000	-0.0403***	0.0468***	-0.0065***
	[0.0017]	[0.0021]	[0.0006]
Married	0.0120***	-0.0300***	0.0179***
	[0.0013]	[0.0016]	[0.0013]
Age	-0.0246***	0.0128***	0.0118***
	[0.0006]	[0.0007]	[0.0002]
Age squared	0.0003***	-0.0002***	-0.0001***
	[8.13e-06]	[9.01e-06]	[2.48e-06]
Immigrant	0.0119**	-0.0270***	0.0151***
	[0.0057]	[0.0054]	[0.0010]
Lived in area 5 yrs. or less	0.0204***	-0.0072***	-0.0132***
	[0.0018]	[0.0015]	[0.0007]
Female	0.0769***	-0.0343***	-0.0426***
	[0.0026]	[0.0026]	[0.0013]
Black, non-Hispanic	0.0181***	0.0124*	-0.0305***
	[0.0051]	[0.0068]	[0.0021]
White, non-Hispanic	-0.0350***	0.0218***	0.0132***
	[0.0047]	[0.0069]	[0.0027]
Hispanic	-0.0006	0.0206***	-0.0200***
	[0.0039]	[0.0052]	[0.0024]
High school education	-0.0530***	0.0529***	0.0001
	[0.0014]	[0.0018]	[0.0008]
Some college education	-0.0667***	0.0625***	0.0042***
	[0.0019]	[0.0023]	[0.0009]
College graduate	-0.0826***	0.0711***	0.0116***
	[0.0023]	[0.0031]	[0.0017]
Own home	-0.0007	-0.0100***	0.0107***
	[0.0010]	[0.0016]	[0.0011]
Census Tract Controls			

	Weak Network		
Variables	Not Working	Paid Employment	Self-Employed
Unemployment Rate	0.5460***	-0.4953***	-0.0506***
	[0.0276]	[0.0326]	[0.0180]
Median Household Income	-1.65e-07***	-3.82e-07***	5.47e-07***
	[3.89e-08]	[4.26e-08]	[1.86e-08]
Population Density	0.0008***	-0.0011***	0.0003***
	[0.0002]	[0.0002]	[0.0001]

Note: Regressions include occupation and MSA fixed effects.
Robust standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table V: Multinomial Logit- Average Marginal Effects; Strong Network

	Strong Network		
Variables	Not Working	Paid Employment	Self-Employed
Social Capital			
High Individual Social Capital	0.0864***	-0.0756***	-0.0109***
	[0.0050]	[0.0047]	[0.0009]
High Community Social Capital	-0.0214***	0.0123	0.0091***
	[0.0067]	[0.0076]	[0.0032]
Individual Characteristics			
Household total income GE $30,000	-0.0536***	0.0577***	-0.0041***
	[0.0025]	[0.0027]	[0.0004]
Married	0.0152***	-0.0328***	0.0176***
	[0.0013]	[0.0015]	[0.0012]
Age	-0.0173***	0.0065***	0.0108***
	[0.0005]	[0.0006]	[0.0003]
Age squared	0.0002***	-0.0001***	-0.0001***
	[6.25e-06]	[7.40e-06]	[3.00e-06]
Immigrant	0.0184***	-0.0325***	0.0141***
	[0.0058]	[0.0054]	[0.0010]
Lived in area 5 yrs. or less	0.0341***	-0.0185***	-0.0155***
	[0.0020]	[0.0018]	[0.0006]
Female	0.0724***	-0.0305***	-0.0419***
	[0.0025]	[0.0026]	[0.0012]
Black, non-Hispanic	0.0164***	0.0125*	-0.0289***
	[0.0051]	[0.0069]	[0.0023]
White, non-Hispanic	-0.0584***	0.0423***	0.0162***
	[0.0055]	[0.0076]	[0.0026]
Hispanic	-0.0024	0.0213***	-0.0189***
	[0.0040]	[0.0051]	[0.0026]
High school education	-0.0555***	0.0558***	-0.0003
	[0.0020]	[0.0022]	[0.0008]
Some college education	-0.0835***	0.0764***	0.0071***
	[0.0030]	[0.0031]	[0.0010]
College graduate	-0.1039***	0.0862***	0.0176***
	[0.0026]	[0.0028]	[0.0014]
Own home	0.0012*	-0.0115***	0.0103***
	[0.0007]	[0.0015]	[0.0010]
Census Tract Controls			

	Strong Network		
Variables	Not Working	Paid Employment	Self-Employed
Unemployment Rate	0.5471***	-0.5012***	-0.0459**
	[0.0277]	[0.0329]	[0.0184]
Median Household Income	-1.59e-07***	-3.92e-07***	5.51e-07***
	[4.61e-08]	[4.94e-08]	[1.80e-08]
Population Density	0.0006***	-0.0010***	0.0004***
	[0.0002]	[0.0002]	[0.0001]

Note: Regressions include occupation and MSA fixed effects.
Robust standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table VI: Immigration and Social Capital

	Immigrants			Natives		
Variables	Not Working	Paid Employment	Self-Employment	Not Working	Paid Employment	Self-Employment
Trust						
High Individual	-0.0015	0.0023	-0.0008	0.0202***	-0.0162***	-0.0041***
	[0.0056]	[0.0039]	[0.0024]	[0.0036]	[0.0034]	[0.0008]
High Community	0.0013	0.0007	-0.0019	-0.0172***	0.0143**	0.0029
	[0.0059]	[0.0060]	[0.0023]	[0.0047]	[0.0059]	[0.0029]
Weak Network						
High Individual	-0.0449***	0.0316***	0.0133***	-0.0064***	0.0022	0.0042***
	[0.0031]	[0.0030]	[0.0014]	[0.0017]	[0.0019]	[0.0009]
High Community	0.0007	-0.0001	-0.0006	-0.0052	-0.0008	0.0060***
	[0.0053]	[0.0056]	[0.0053]	[0.0046]	[0.0059]	[0.0020]
Strong Network						
High Individual	0.0979***	-0.0823***	-0.0156***	0.0929***	-0.0814***	-0.0115***
	[0.0024]	[0.0026]	[0.0012]	[0.0048]	[0.0048]	[0.0009]
High Community	0.0042	-0.0004	-0.0038	-0.0223***	0.0128*	0.0095***
	[0.0087]	[0.0075]	[0.0043]	[0.0066]	[0.0077]	[0.0032]
Observations	N=2,488,000			N=18,070,000		

*Each measure of social capital is from a separate regression. Robust standard errors are clustered at the state level. Each model contains the same controls detailed in Table III but are not repeated here for brevity.

*** p<0.01, ** p<0.05, * p<0.1

Table VII: Urban, Rural Status and Social Capital

	Urban			Rural		
Variables	Not Working	Paid Employment	Self-Employment	Not Working	Paid Employment	Self-Employment
Trust						
High Individual	0.0196***	-0.0164***	-0.0032***	-0.0081	0.0098*	-0.0018
	[0.0029]	[0.0028]	[0.0006]	[0.0059]	[0.0058]	[0.0029]
High Community	-0.0128***	0.0108***	0.0021	-0.0085	0.0015	0.0071
	[0.0028]	[0.0033]	[0.0017]	[0.0090]	[0.0094]	[0.0048]
Weak Network						
High Individual	-0.0185***	0.0125***	0.0060***	0.0092*	-0.0194***	0.0102***
	[0.0015]	[0.0018]	[0.0012]	[0.0049]	[0.0053]	[0.0025]
High Community	-0.0024	-0.003	0.0054***	-0.0084	-0.0003	0.0087***
	[0.0037]	[0.0049]	[0.0014]	[0.0060]	[0.0071]	[0.0031]
Strong Network						
High Individual	0.0877***	-0.0766***	-0.0111***	0.0959***	-0.0840***	-0.0119***
	[0.0057]	[0.0052]	[0.0011]	[0.0059]	[0.0056]	[0.0013]
High Community	-0.0085	0.0053	0.0032	-0.0272***	0.0116	0.0156***
	[0.0052]	[0.0047]	[0.0026]	[0.0079]	[0.0096]	[0.0042]
Observations	N=15,050,000			N=1,909,000		

*Each measure of social capital is from a separate regression. Robust standard errors are clustered at the state level. Each model contains the same controls detailed in Table III but are not repeated here for brevity.

*** p<0.01, ** p<0.05, * p<0.1

Table VIII: Gender and Social Capital

	Females			Males		
Variables	Not Working	Paid Employment	Self-Employment	Not Working	Paid Employment	Self-Employment
Trust						
High Individual	0.0251***	-0.0251***	-0.0001	-0.0042	0.0078**	-0.0036***
	[0.0037]	[0.0037]	[0.0006]	[0.0032]	[0.0031]	[0.0011]
High Community	-0.0289***	0.0258***	0.003	-0.0024	0.0011	0.0013
	[0.0063]	[0.0070]	[0.0023]	[0.0047]	[0.0064]	[0.0035]
Weak Network						
High Individual	-0.0012	-0.0034	0.0046***	-0.0269***	0.0180***	0.0089***
	[0.0027]	[0.0023]	[0.0010]	[0.0017]	[0.0021]	[0.0015]
High Community	-0.0093**	0.0033	0.0060***	0.0003	-0.0049	0.0046***
	[0.0041]	[0.0047]	[0.0012]	[0.0053]	[0.0066]	[0.0017]
Strong Network						
High Individual	0.1166***	-0.1069***	-0.0098***	0.0464***	-0.0353***	-0.0111***
	[0.0076]	[0.0071]	[0.0010]	[0.0028]	[0.0032]	[0.0011]
High Community	-0.0383***	0.0293***	0.0090***	-0.0039	-0.0052	0.0092**
	[0.0089]	[0.0100]	[0.0025]	[0.0056]	[0.0064]	[0.0041]
Observations	N=10,260,000			N=10,290,000		

*Each measure of social capital is from a separate regression. Robust standard errors are clustered at the state level. Each model contains the same controls detailed in Table III but are not repeated here for brevity.

*** p<0.01, ** p<0.05, * p<0.1

Table IX: Education and Social Capital

	College Graduate			Non-College Graduate		
Variables	Not Working	Paid Employment	Self-Employment	Not Working	Paid Employment	Self-Employment
Trust						
High Individual	0.0032	0.0008	-0.0041**	0.0176***	-0.0148***	-0.0028**
	[0.0032]	[0.0027]	[0.0018]	[0.0037]	[0.0035]	[0.0011]
High Community	-0.0064*	0.0073	-0.0009	-0.0181***	0.0152**	0.0028
	[0.0037]	[0.0062]	[0.0038]	[0.0053]	[0.0060]	[0.0025]
Weak Network						
High Individual	-0.0257***	0.0156***	0.0101***	-0.0093***	0.0067***	0.0026***
	[0.0020]	[0.0018]	[0.0012]	[0.0021]	[0.0022]	[0.0009]
High Community	0.0030**	-0.0104***	0.0075**	-0.0062	0.0018	0.0043***
	[0.0014]	[0.0040]	[0.0035]	[0.0050]	[0.0055]	[0.0008]
Strong Network						
High Individual	0.1017***	-0.0884***	-0.0133***	0.0865***	-0.0753***	-0.0113***
	[0.0083]	[0.0068]	[0.0023]	[0.0040]	[0.0042]	[0.0007]
High Community	-0.0208***	0.0134*	0.0075*	-0.0216***	0.0127	0.0089***
	[0.0053]	[0.0079]	[0.0043]	[0.0075]	[0.0078]	[0.0029]
Observations	N=4,963,000			N=15,600,000		

*Each measure of social capital is from a separate regression. Robust standard errors are clustered at the state level. Each model contains the same controls detailed in Table III but are not repeated here for brevity.

*** p<0.01, ** p<0.05, * p<0.1

Chapter II: Amenities and the College-Educated: A Gentrification Perspective

I. Introduction

Gentrification is an important topic within public policy, being a subject of debate and interest for economists, city planners, and politicians. Its potential for displacement of incumbent residents has led many to label this urban phenomenon as a problem. While there is no set definition of gentrification, it is considered a process by which highly educated, wealthier individuals locate in previously low-income, high-minority neighborhoods in urban areas. Neighborhoods experiencing significant increases in educational attainment or income are considered gentrifying in post 2000 metropolitans. Often with gentrification, neighborhoods gain access to establishments that were previously lacking, namely consumption amenities such as bars, restaurants, and boutiques. In general, cities want to increase the number of highly educated individuals as increases in college-educated adults is associated with increased wages of lower-skilled workers as well as promoting economic growth (Moretti 2004; Pink-Harper 2015).Therefore, gentrification may not be a problem but a policy tool that city governments actively pursue. Gentrification and the mechanisms through which it is perpetrated need to be understood in order to formulate appropriate policies. One of these potential mechanisms is the establishment location of consumption amenities.

The purpose of this paper is the development of a model of the gentrification process whereby gentrification and amenity establishment locational choice influence each other. I utilize three-sample three-stage least squares to address the colocation of gentrification, defined as growth in a location's college-educated population and consumption amenities. At its essence, the relationship between the two describes their tendency to co-exist in the same location. In this way, the model is almost a supply and demand model for space.

Currently, there are two main theories on the causal mechanisms of the gentrification process: (1) gentrification results from highly educated residents being attracted by growth in urban high paying jobs (Edlund, Machado, and Sviatschi 2019) or (2) gentrification is caused by preferences for urban consumption amenities (Couture and Handbury, 2019). I focus on the second mechanism, where gentrification and increasing consumption amenities are modeled as happening concurrently – thereby influencing each other. This conceptual model is theoretically similar to a "supply and demand" model: the market is for consumption amenities, firms are the suppliers, and the highly skilled are the demanders.

It is important to understand that there is a discourse between the historical definitions of gentrification and the operational definitions employed by researchers. Historically, gentrification is the process by which low-income minorities are displaced from their neighborhoods by wealthy, highly educated, predominantly white new residents. Displacement is thought to potentially occur because property values increase with an influx of more affluent families. Incumbent renters and homeowners will see their cost rise (through increases in rents and property taxes) and be forced to move. Many studies from the past 20 years, however, find that displacement is not extensively occurring in neighborhoods as they undergo the process of gentrification (Ding and Hwang, 2016; Martin and Beck, 2018; McKinnish et al., 2010; Vigdor et al., 2002). Operationally, gentrification is typically defined as increases in a neighborhood's place in the distribution of income, the share of college graduates, or housing value (Kolko, 2007; Lester and Hartley, 2014; Meltzer and Ghorbani, 2017). This paper will make use of this operational definition of gentrification while not addressing issues of displacement. Neighborhoods are considered gentrified if the neighborhood experiences an increase in the share of college-educated adults over a relatively short period of time since the start of the 2000s

(Brummet and Reed, 2019). I utilize both binary and continuous measures of gentrification. Gentrification is constructed with an underlying measurement of education attainment to emphasize *who* are the people behind gentrification as opposed to a symptom of gentrification (higher income and higher home prices). Consumption amenities are narrowly defined used six-digit NAICS codes.[12]

This paper contributes to the literature in three ways. First, the main contribution of this paper is to model gentrification in an environment with reverse causality with consumption amenities. Previous literature (e.g., Couture and Handbury, 2019 and Baum-Snow and Hartley 2020) serve as the precursor to my study by showing there are strong correlations between gentrification and consumption amenities. Empirically estimating the causal mechanism of gentrification and amenity establishments allows one to have a more complete understanding of the process of gentrification. Secondly, in contrast to other studies that typically focus on gentrification taking place in poor neighborhoods, I focus my analyses not only on neighborhoods that are extremely poor or have low educational attainment, but also consider more middle-income/middle-educated neighborhoods. Studying gentrification in the Chicago area, Hwang and Sampson (2014) find that neighborhoods at the lower levels of the income distribution stay consistently poor and the incomes in the upper levels of the income distribution are consistently rich. The neighborhoods in the middle of the income distribution can go either direction. A restricted measure of gentrification that only looks at the poorest neighborhoods (or ones with the lowest levels of human capital) during the initial period would leave these middle distribution neighborhoods out of the analysis.

[12] See Appendix B for more details.

Finally, this paper adds to the current understanding of the potential positive benefits of neighborhood gentrification. The production of consumption amenities (e.g., the workers of consumption amenities) often require low skill. Therefore, gentrification may lead to more job opportunities for incumbent low-skill residents. If policymakers are truly concerned about displacement, then incentivizing these firms to hire local workers with low educational attainment could serve as an important policy objective.

The rest of this paper is as follows: Section II synthesizes in more detail the literature concerning gentrification over the past several decades; Section III describes the data that is utilized in this paper; Section IV describes the empirical methodology; Section V presents results, and finally Section VI concludes.

II. Literature Review

Early gentrification literature focuses primarily on the years between 1970 and 2000 (Henig 1980; Vigdor, Massey, and Rivlin 2002) . Because of the influx of wealthy residents, gentrification has been characterized in the media by its displacement effects of poor, incumbent residents. These effects usually emerge through changes in housing prices. While wealth transfers to homeowners in gentrifying neighborhoods seems like a positive effect, it can also be harmful to incumbent residents. Neighborhoods and communities that don't have protections on property tax increases may see their incumbent residents forced out. Additionally, high home values translate to higher rents, and large increases in rents can hurt incumbent renters.

Studies conducted on data from the 1990s find mobility rates of poor residents in gentrifying neighborhoods not differing drastically from mobility rates of the same type of residents in non-gentrifying neighborhoods (McKinnish, Walsh, and White 2010; Vigdor, Massey, and Rivlin 2002). Post-2000 gentrification work has not found much evidence of residential displacement (Ding et al., 2016; Martin and Beck, 2018). All the works mentioned in this section use an operational definition of gentrification (i.e., the neighborhoods economic standing increased over a time period). While gentrification is often thought of synonymously with displacement, the empirical literature presented here finds few associations with operationally defined gentrification and displacement.

While often viewed negatively because of presumed displacement effects, gentrification could lead to beneficial results for incumbent residents. One of the positive aspects of gentrification could be the potential for increased employment opportunities for incumbents, as gentrification is often associated with an influx of new wealthy households and business into previously poor neighborhoods with high unemployment. This is particularly important if the residents are experiencing spatial mismatch. Many of the cities analyzed in the gentrification literature have at some point experienced spatial mismatch of their central city's African-American population (Hellerstein, Neumark, and McInerney 2008; Holzer 1991; Kain 1968). In contrast to the lack of displacement effects in earlier work, there is some evidence that gentrification augments spatial mismatch, with neighborhoods experiencing gentrification seeing local resident job losses, and therefore, worsening spatial mismatch (Meltzer and Ghorbani, 2017). While there are more jobs in gentrifying neighborhoods, these jobs are not going to local residents. The aforementioned piece, however, was for only one MSA, New York. However, this potential negative implication is polarized by evidence of gentrifying neighborhoods benefitting incumbents financially in the form of better credit scores and lower default rates (Hartley 2013). Empirically, Ding and Hwang (2016) find that gentrification is associated with increases in vulnerable populations' credit scores among residents who remain in gentrifying neighborhoods.

These studies suggest the potential for better economic opportunities of incumbent residents, perhaps in the form of jobs.

While earlier works concerning gentrification examine whether incumbent residents are harmed by the influx of new rich inhabitants, there is also a shifting focus to identifying the drivers behind recent gentrification. Baum-Snow and Hartley (2020) and Couture and Handbury (2019) provide comprehensive evidence of neighborhood changes in almost all major U.S. cities. The latter finds that initial levels of consumption amenities have more power to explain urbanization than changes in the relative availability of jobs. Additionally, increasing taste for non-tradables is the biggest contributing factor towards the increase in the young college educated shares of downtown Census Tracts. One of the most important factors for young, college-educated city dwellers is the quality of the amenities.

Similarly, Baum-Snow and Hartley (2020) find that changes in amenity valuation after 2000 encourage college-educated whites to move to central city neighborhoods. Less educated whites appear to remain in central cities, but less educated minorities are displaced from central city neighborhoods. The authors' results support previous findings that urban revitalization is primarily found very close to the central business district, usually within 2-3 kilometers. However, none of these studies examine how gentrification has effect changes in amenities. Conceptually modeling gentrification and amenities simultaneously shows how the process of gentrification can be characterized by both a growth in college-educated residents and growth in consumption amenities.

From a business perspective, firms are attracted to gentrifying areas. For instance, Lester and Hartley (2014) find that gentrification is associated with a mildly positive impact on the overall number of jobs. The big take away from their results lies in the type of jobs that are coming to gentrifying neighborhoods. In neighborhoods undergoing revitalization, restaurants and retail jobs tend to replace manufacturing jobs. Similarly, Baum-Snow and Hartley (2020) find that new businesses are not likely to be displaced by gentrification with "[g]entrifying neighborhoods more likely to attract new types of services than non-gentrifying."

Glaeser et al. (2018) use Yelp data to quantify neighborhood change typically associated with gentrification. Operationalized gentrification as measured by changes in the college-educated has some of the highest correlations with types of establishments that could proxy "amenities" such as grocery stores, cafes, and bars. This also calls to attention how businesses, not consumers, can contribute to gentrification. Recent findings by Su (2018) suggest that while the change in value of time is an initial force behind gentrification, its effects are substantially magnified by endogenous amenity improvement. He calls for, "future research [speaking] to how firms' location decision responds to workers geographic sorting" -- this is the focus of this paper.

Outside of academics, the potential benefits of gentrification are starting to gain traction. Popular press articles, such as a recent article published in *The Econom*ist, have now started to call attention to the idea that gentrification is many more things than just displacement. While gentrification is often looked at with disdain, there are potential benefits when an area undergoes gentrification such as access to more amenities, as well as more racial and economic integration.[13]

III. Data

This paper utilizes both publicly available data to measure gentrification and private data to provide detailed information on consumption amenity establishments. Data pertaining to

[13]The Economist (2018); Link to source https://www.economist.com/united-states/2018/06/21/in-praise-of-gentrification.

gentrification is obtained via National Historical Geographic Information Systems via IPUMS.org (Manson et al. 2019).[14] These data contain educational attainment of individuals between 25-44 as well as neighborhood median income, where neighborhood is defined as a Census Tract. I utilize tables from the 2000 U.S. Decennial Census, the 2005-2009 5-Year American Community Survey (ACS) and the 2012-2016 5-Year ACS. With these three data sources I have snap shots of the neighborhoods for 3 non-overlapping years. For the two 5-Year average tables, I associate the midpoint, 2007 and 2014, as the years with which I match these data to the annual establishment level data. Both the 2000 U.S. Decennial Census and the 2005-2009 5YR ACS are tabulated to 2000 Census Tract geographies. In order for these to be comparable to the 2012-2016 5Yr ACS, all data from the two samples are cross walked to 2010 Census tract geographies using a crosswalk table from U.S. Census Bureau (2020).

Establishment level data comes from ReferenceUSA database.[15] This is a proprietary database that is accessible via the Georgia State University library for the years 1997-2017. Each establishment has a six-digit industry code. These detailed codes are called North American Industry Classification System (NAICS). Outside of restricted data centers, it is rare to find establishment level data with this detailed industry codes with fine geographic information. Publicly available establishment data, such as the Statistics of U.S. Businesses (SUSB), does contained detailed industry information for establishments across the U.S., but the narrowest geography available is state-level.[16] For this reason, Reference USA data is preferred to other

[14] Manson, S., Schroeder, J., Van Riper, D., & Ruggles, S. (2019); For more information please see: https://nhgis.org/

[15] ReferenceUSA, 2020; For more information on this database please follow the link: http://www.referenceusa.com.eu1.proxy.openathens.net/Home/Home

[16] U.S. Census Bureau (2020); For more information on the SUSB, please follow the link: https://www.census.gov/programs-surveys/susb.html

data sources. Additionally, this data contains additional information on establishment age as well as sales volume. For this paper, only the years 2000, 2007, and 2014 are used as they best match the publicly available data used. Detailed industry codes are needed in order to classify an industry as a consumption amenity as well as to attempt to proxy for amenity quality. For example, a high-end grocery store, such as a Whole Foods [NAICS 445110], has a separate firm identifier than a Seven-Eleven [NAICS 445120].[17] In publicly available data, only broad industry codes are listed, and these two types of establishments are of the same broad industry classification [4451]. These data are at the establishment level but are coded to 2010 Census Tracts. Individual establishment level data is collapsed to the Census Tract level to obtain Census Tract counts of establishments in each 6-digit industry-type. Table B-I in the appendix lists the industry codes used to classify establishments as consumption amenities. This classification is motivated by establishments and firms found to be prominent in gentrifying areas as detailed in Couture and Handbury (2019) and Glaeser et al. (2018).

Census Tract share of college graduates is used to operationalize my gentrification measure. The focus on college graduates, as opposed to median income or median housing value, allows us to focus on *who* the gentrifiers are as opposed to *how* the gentrifiers influence the neighborhood change. More so, it is the highly educated who have been found to be most associated with increased presence of consumption amenities. This work examines a spectrum of gentrification measures both continuous and discrete.

Finally, I restrict all analyses to neighborhoods in metros with populations over 1,000,000 in 2000. This restriction keeps my samples similar to the samples used in the existing

[17] NAICS Association (2018); For more information on NAICS please follow the link: https://www.naics.com/search/

gentrification literature (Couture and Handbury, 2019; Glaeser et al., 2018; Meltzer and Ghorbani, 2017). A list of these metros is available in Table X. Approximately 27,000 Census tracts are included in my sample of these 50 metros. This constitutes slightly over a third of Census tracts in the 2010 delineations (U.S. Census Bureau, 2020).

IV. Methodology

This section will describe the methodology used to define gentrification and amenities as well as how the two will be modelled together. The time span of my study is between 2007 and 2014.

IV.A. Gentrification

Gentrifiers are typically young, educated, and wealthy adults (Baum-Snow & Hartley, 2020; Couture & Handbury, 2019; Meltzer & Ghorbani, 2017; Su, 2018). Following Brummet and Reed (2019), gentrification in this paper is measured by the share of college graduates in a neighborhood. This measure is consistent with the idea that urban revival (another phrase used to describe gentrification) is driven by this group across racial categories, except for college educated Blacks where the growth in this demographic is still concentrated in suburbs (Couture and Handbury 2019). I construct several measures of gentrification, both dichotomous and continuous, in order to compare my results with the previous literature's findings. The first measure of gentrification is an indicator for whether a neighborhood has gentrified. A neighborhood is considered gentrified if the share of college graduates (*SCG*) in census tract *t* in Core-based Statistical Area (CBSA) *c* was in the bottom three quintiles of the CBSA's distribution of share of college graduates in 2007 and in the top two quintiles (or top quintile) of the CBSA's distribution of share of college graduates in 2014. This restriction is done to create a group of neighborhoods that have a common set of circumstances (low-to-moderate-education) during an initial time period and therefore, have the ability to move most upwards in the distribution. However, one drawback is that it does treat small movers and big movers within the CBSA distribution the same. This indicator measure of gentrification is described in equation (1) below.

$$\begin{cases} Gentrified = 1 \; if \;\; SCG_{i,c,2007} \leq SCG_{i,c,2007}^{60th} \; \boldsymbol{and} \; SCG_{i,c,2014} \geq SCG_{i,c,2014}^{60th} \\ Gentrified = 0 \; otherwise \end{cases} \tag{1}$$

I compare this measure of gentrification with a neighborhood's (census tract i) change in college graduates, ages 25-44, between 2007 and 2014 in Core-Based Statistical Area c relative to the total number of 25-44-year old's in 2007. This very similar to the measure utilized in Brummet and Reed (2019) . It captures the change in college graduates, ages 25-44, relative to the initial population, ages 25-44, in CBSA c in Census Tract i. The change in college graduates is weighted by the initial population as to not simultaneously correlate a change in college graduates during this time period with people of less educational attainment moving away and people with more educational attainment moving in.

$$\Delta CG_{c,i} = \frac{CG_{i,c,2014} - CG_{i,c,2007}}{Total\ Population^{25-44}{}_{i,c,2007}} \tag{2}$$

I utilize the measure from (2) on the sample of neighborhoods considered by the dichotomous measure described in equation (1) as well as in a specification where there is no restriction on initial neighborhood conditions (i.e., the share of college graduates). Equation (2) is more flexible than the dichotomous definition of gentrification since it does not arbitrarily assign a label of "gentrified" to any neighborhood, but simply compares the growth in a key characteristic used to identify the process of gentrification.

IV.B. Amenities

From the ReferenceUSA data, establishments are mapped to Census tracts. Establishments are considered an amenity or not based on their six-digit NAICS code. Utilizing detailed industry codes allows me to proxy for amenity quality. I measure growth in amenities as the change in amenity establishments (AE) in CBSA c in Census Tract i, between 2007 and 2014, relative to the number of total establishments in CBSA c in Census Tract i, in 2007. The explicit measure of amenities utilized in the empirical model is detailed below:

$$\Delta AE_{i,c} = \frac{AE_{i,c,2014} - AE_{i,c,2007}}{Total\ Establishments_{i,c,2007}} \quad (3)$$

The change in amenity establishments is also weighted by 2007 level establishments for a similar reason why the change in college graduates is weighted by its 2007 level of college graduates: to not systematically correlate the measurement with establishments moving out and being replaced by other establishments moving into the neighborhood.

IV.C. Simultaneous Determination of Gentrification and Amenities

The empirical contribution of this paper is estimating the causal relationship between amenities and gentrification using three-stage least squares (3SLS) methodology. This method was pioneered by Zellner and Theil (1962) and "[uses] the two-stage least squares estimated moment matrix of the structural disturbances to estimate all coefficients of the entire system simultaneously." 3SLS allows for the error terms across equations to be correlated. Therefore, previous work such as Baum-Snow and Hartley (2020) and Couture and Handbury, (2019) fail to account for this cross correlation in explaining the relationship between consumption amenities and gentrification, operationalized by changes in a neighborhood's educational attainment. Accounting for this cross correlation of the system errors terms is the main advantage of the 3SLS methodology over traditional two stage least squares.

The estimating equations are defined as follows:

$$\left(Gent_{c,i} = 1\right) = \alpha + \beta_1 \%\Delta AE_{c,i} + \beta_2 SCG_{c,i_{2000}} + X_{c,i} + \gamma_c + \varepsilon_{c,i} \quad (4)$$

$$\Delta AE_{c,i} = \mu + \omega_1\left(Gent_{c,i} = 1\right) + \omega_2\, SEA_{c,i_{2000}} + Y_{c,i} + \lambda_c + \sigma_{c,i} \quad (5)$$

Equations (4) and (5) are estimated via three-stage least squares (3SLS). The final stage of the estimating procedure is done via generalized least squares (GLS).[18] Therefore, the results from estimating equations (4) and (5) should be assumed as coming from GLS. In equations (4) and (5) γ_c and λ_c are CBSA fixed effects. $X_{c,i}$ are Census Tract percentage change in the number of renters and the percentage change in unemployment from 2007 to 2014. $Y_{c,i}$ are Census Tract percentage change in unemployment from 2007 and 2014 and median establishment age.[19] $SCG_{c,i_{2000}}$ denotes the neighborhood's initial share of college graduates in 2000. $SEA_{c,i_{2000}}$ denotes the neighborhood's initial share of amenity establishments in 2000.

In each of the equations an exogenous regressor identifies the equations. All exogenous regressors act as instrument for both the endogenous variables in the system in the first stage.[20] The estimation of this system is different than pure instrumental variables. In an instrumental variable system, one would directly use an instrument Z for the endogenous variables. Here the exogenous variables are part of the set of simultaneous equations. Typical for instruments, the exclusion restriction requires that each exogenous variable only influences the model through the endogenous variable. For gentrification, initial shares of college graduates (in 2000) will determine how the share of college graduates over 2007 and 2014 change but shouldn't influence

[18] ΔCG is replaces $Gent = 1$ when the gentrification measurement is a continuous one.
[19] Median establishment age is included in the amenity equation due the relationship between firm growth and age of existing establishment age found in Faberman (2011)
[20] I utilize the *reg3* command in Stata to estimate the system of equations detailed in (4) and (5).

the current percent change in amenities. Previous work has found that urban revival pre-2000, was more focused on individuals desire to live closer to work and more recent urban revival is due in part to individuals' desire to locate near consumption amenities (Couture and Handbury, 2019). Its effect in the system is measures by β_2 in equation (4). A similar argument is used for the initial (2000) share of amenities and current gentrification of a census tract. It's influence in the system of equations is measured by ω_2 in equation (5). In all analyses, census tracts are weighted by their 2000-level population.

Table X displays means of ΔCG and ΔAE for the sample that is restricted to only neighborhoods whose 2007 share of college graduates was in the bottom 60th percentile of the CBSA distribution and the correlation between the two. There are approximately 16,000 neighborhoods in this sample. Orlando-Kissimmee-Sanford, FL has the strongest correlation between gentrification and consumption amenities at 0.2684.

V. Results

I first briefly discuss results from the first stage in the 3SLS setup. Then the final stage estimates compare the dichotomous measure of gentrification (equation 2) to the continuous measure of gentrification described in equation (3), for the full sample and then for the restricted sample.

V.A. First Stage Results

Table B-II and B-III present the results for the first stage of the 3SLS. Table B-II are the first stage results from the specification where the sample of neighborhoods is restricted to only neighborhoods in the bottom 60% percentile of the CBSA distribution of college graduate shares in 2007. The gentrification measures are described by equations (1) and (2). Table B-III does not restrict the sample and the measure of gentrification is measured first as the $\%\Delta SC$ detailed in equation (2) and then as an indicator if a neighborhood's $\%\Delta SG$ is in the top 40% of the CBSA distribution of $\%\Delta CG$.

In 3SLS, all exogenous regressors enter the first stage. For each equation in the system, a good identification strategy needs variables that identify the system. For the gentrification equation (7), the unique regressors are 2000 share of college graduates ages 25-44 and the percentage change in renters. For the amenity equation, the median establishment age and the 2000 share of amenities are the identifying regressors.

Table B-II restricts the neighborhoods to the set of neighborhoods with college education shares in the bottom 60% of the CBSA distribution of college-educated shares. The 2000 shares of college graduates strongly predict both the propensity for a neighborhood to gentrify and increases in $\%\Delta AE$. Table B-III does not restrict the neighborhoods by their 2007 share of college graduates. The binary gentrification measure detailed in equation (2) is 1 if the ΔCG is in the top 40% of the CBSA distribution of ΔCG. Overall, in both equations, the exogenous identifiers are statistically significant in predicting the endogenous regressors.

V.B. 3SLS: Initial Condition Restrictions Gentrification

As was mentioned earlier, 3SLS not only deals with the endogeneity between gentrification and consumption amenities but allows for there to be cross correlation in the disturbance terms of the system of equations ($\varepsilon_{c,i}$ and $\sigma_{c,i}$ in equations (4) and (5)). I first look at the simultaneous relationship between gentrification and amenity establishments by examining neighborhoods who share of college graduates was in the lower 60th percentile of its CBSA distribution. This sample selection is similar to sample selection in previous gentrification literature (Martin and Beck, 2018; McKinnish et al., 2010; Meltzer and Ghorbani, 2017). In contrast to previous works, I make this restriction in order to capture not simply potentially poor neighborhoods but also more middle-income neighborhoods as well.[21]

Table XI presents the results from three specifications on this restricted sample. The first specification is detailed in columns (1) and (2). Here gentrification is a dichotomous measure equal to 1 if the neighborhood moves up in its CBSA distribution of share of college graduates to the top 40th percentile in 2014. When controlling for the potential simultaneity between both endogenous variables, I fail to find conclusive evidence that increases in amenities measured by growth in establishments relative to total establishments in 2007 increase the likelihood a neighborhood gentrifies significantly in 2014. Columns (3) and (4) are results from a more "restrictive" gentrification measure, where a neighborhood is labelled as gentrified if the neighborhood moves to the upper 20th percentile in the 2014 CBSA distribution of share of college graduates. In both dichotomous gentrification specifications, neighborhoods that gentrify see increases in the number of amenities, as measured by number of establishments, with neighborhoods whose increase in their share of college graduates seeing increases by 0.01592 percentage points (1.592/100). Compared to the mean increase in $\%\Delta AE$ of gentrifying neighborhoods of 0.36, gentrification leads to approximately 4% increase in $\%\Delta AE$. Columns (5) and (6) utilize the continuous gentrification measure described in equation (2) but still on the initial condition restricted sample. Compared to the dichotomous system, where big movers and small movers within the CBSA distribution are treated the same, here the actual growth or change in the share of college graduates is modelled directly in the system. The results are

[21] I use the term "middle-income", but I still only use measures of educational attainment to make gentrification determinants.

consistent with the previous two specification albeit with a slightly poorer fit in the amenity establishment equation.

V.C. 3SLS: No Initial Condition Restrictions Gentrification

I next examine the simultaneous nature between gentrification and amenity establishments utilizing my entire sample of neighborhoods. Unlike the specification detailed in Table II, the specifications detailed in Table XII put no restrictions on a neighborhood's eligibility to gentrify. Therefore, middle- and high-income neighborhoods are included. The underlying measure of gentrification is that of equation (2).[22] A neighborhood simply gentrifies if it's ΔCG is in the upper 40th (or 20th) of the CBSA distribution of ΔCG. Therefore, the binary indicator of gentrification as seen in Table XII, is based off a neighborhood's place in the ΔCG distribution. These specifications are detailed in columns (1)-(4) in Table XII. Neighborhoods that see growth in their college graduates see increases in their amenities (0.002 and 0.005). There is again no conclusive evidence that increases in amenities is driving gentrification. Finally, columns (5) and (6), simply let the ΔCG enter as the measure of gentrification. While the results here are consistent with all previous specifications, the fit is exceedingly poor (a negative adjusted R^2). These results indicate the importance in the gentrification literature of identifying the correct counterfactuals. In other words, the initial conditions a neighborhood (here

[22] Specifications run with gentrification specified by equations (1) and (2) were a poor fit.

neighborhoods with initially low shares of college graduates), are the neighborhoods to consider when studying gentrification as operationalized here.

V.D. Falsification-Always High-Education Neighborhoods vs. Always Low-Education Neighborhoods

The final specifications I estimate are akin to "falsification" specifications. I first estimate a model where I only consider tracts in the top 40th percentile of the CBSA distribution of share of college graduates. Then, I indicate a neighborhood as gentrified if that neighborhood remains in the top 40th percentile. I estimate this system to test the hypothesis that the neighborhoods that are important in the discussion of gentrification and amenities are ones that initially start out lower in the CBSA educational attainment distribution. The results from this specification are detailed in Table XIII columns (1) and (2). Here, there is some evidence that growth in amenities increases the likelihood a neighborhood remains in the top 40th percentile. However, neighborhoods that remain in the top 40th percentile (i.e., gentrified), don't offer any evidence of increases the number of amenity establishments.

I next look at only neighborhoods whose share of college graduates was in the bottom 40th percentile of the CBSA distribution in 2007 and indicate "gentrification" if a neighborhood stays there. This is a falsification test because, these neighborhoods always remain educationally poor and therefore do not meet the criterion for gentrification. Neighborhoods that meet this requirement see a decrease in the ΔAE and increases in ΔAE do not indicate any evidence that a neighborhood remains in the lower part of the share of college education distribution. These two specifications shed some light that in terms of growth in amenities, neighborhoods that see increases in college graduates from initially low levels to much higher levels are the neighborhoods that are see the largest increases in the amenities in their neighborhood.

VI. Concluding Remarks

Gentrification is a popular topic of conversation in many cities in the U.S. The attraction of educated adults can have positive spillover affects to less educated adults in the form of higher wages (Moretti 2004). To my knowledge, this paper is the first to put a simultaneous framework on the relationship between gentrification and consumption by modelling the two entities using simultaneous empirical strategy. This is the main contribution of the paper.

I find when controlling for the simultaneous nature between the two entities, gentrification causes a neighborhood to see a growth in its amenities, measured here by amenity establishments. However, there is no conclusive evidence that gentrification is leading to a neighborhood's growth in amenities (as measured by changes in amenity producing establishments). This is in direct contrast to previous literature such as Couture and Handbury (2019). This may be due in part to the failure to model the simultaneous decision of both establishments and individuals co-locating, which the simultaneous model considers.

One key takeaway from this work is that gentrification is not solely about displacement. Many things are occurring when a neighborhood is undergoing gentrification. The results in this paper speak to the potential benefits from gentrification. Gentrification is found to increase neighborhood amenities. Individuals in neighborhoods with higher amenities have residents that report higher levels of happiness than individuals in low amenity and rural areas (Cox and Streeter, 2019). While the economic costs of displacement are no concern to be ignored, it is important to consider other phenomena that are occurring and how that phenomenon affect individuals.

One drawback of the current analyses is the limited panel nature. Establishment level data with such detailed industry codes is difficult to acquire outside of purchasing the data (i.e., the National Establishment of Times Series) or obtaining it through a government lab (i.e., a Federal

Research Data Center). Future work will expand the current model with more updated data utilizing the Census Bureau's Business Registrar.

Table X: Chapter II Summary Statistics

CBSA	Total Tracts	Eligible to Gentrify* Restricted Sample	# Gentrified	Mean $\%\Delta GC_{i,c}$**	Mean $\%\Delta AE_{i,c}$**	Correlations between Gentrification and Amenities
Atlanta-Sandy Springs-Marietta, GA	456	287	26	0.153	0.4589	0.0259
Austin-Round Rock-San Marcos, TX	198	122	17	0.2562	0.346	-0.024
Baltimore-Towson, MD	477	356	32	0.2557	0.2518	0.0284
Birmingham-Hoover, AL	177	130	7	0.1064	0.4908	0.0501
Boston-Cambridge-Quincy, MA-NH	786	486	33	0.2121	0.2433	0.001
Buffalo-Niagara Falls, NY	244	149	21	0.5874	0.3904	0.1912
Charlotte-Gastonia-Rock Hill, NC-SC	143	102	10	0.4781	0.4175	0.1471
Chicago-Joliet-Naperville, IL-IN-WI	1,441	924	64	0.2369	0.2196	0.0226
Cincinnati-Middletown, OH-KY-IN	400	255	33	0.2561	0.3932	0.0445
Cleveland-Elyria-Mentor, OH	550	344	33	0.2121	0.2503	0.085
Columbus, OH	326	218	15	0.1896	0.2085	0.0287

CBSA	Total Tracts	Eligible to Gentrify* Restricted Sample	# Gentrified	Mean $\%\Delta GC_{i,c}$**	Mean $\%\Delta AE_{i,c}$**	Correlations between Gentrification and Amenities
Dallas-Fort Worth-Arlington, TX	831	591	37	0.2275	0.3867	0.0786
Denver-Aurora-Broomfield, CO	376	236	25	0.3394	0.4734	0.0101
Detroit-Warren-Livonia, MI	1,095	719	58	0.1692	0.313	0.1289
Hartford-West Hartford-East Hartford, CT	245	158	23	0.2034	0.1803	-0.0191
Houston-Sugar Land-Baytown, TX	682	456	32	0.4476	0.175	0.0079
Indianapolis-Carmel, IN	252	175	16	0.2111	0.1348	0.0539
Jacksonville, FL	143	113	9	0.1359	0.2918	0.0035
Kansas City, MO-KS	374	244	19	0.169	0.2897	0.054
Las Vegas-Paradise, NV	234	193	28	0.1814	0.1596	-0.0482
Los Angeles-Long Beach-Santa Ana, CA	2,117	1,340	94	0.2026	0.2099	0.0634
Louisville/Jefferson County, KY-IN	212	138	12	0.1761	0.276	0.0567
Memphis, TN-MS-AR	202	129	11	0.1417	0.1184	0.0124

CBSA	Total Tracts	Eligible to Gentrify* Restricted Sample	# Gentrified	Mean $\%\Delta GC_{i,c}$**	Mean $\%\Delta AE_{i,c}$**	Correlations between Gentrification and Amenities
Miami-Fort Lauderdale-Pompano Beach, FL	570	393	42	0.2126	0.5005	0.0106
Milwaukee-Waukesha-West Allis, WI	326	221	18	0.225	0.0705	0.0675
Minneapolis-St. Paul-Bloomington, MN-WI	655	414	41	0.2066	0.4783	0.0784
Nashville-Davidson--Murfreesboro--Franklin, TN	173	116	16	0.2558	0.2644	-0.0416
New Orleans-Metairie-Kenner, LA	282	204	32	0.4644	0.7733	-0.0255
New York-Northern New Jersey-Long Island, NY-NJ-PA	3,680	2,376	257	0.2731	0.1841	-0.0062
Oklahoma City, OK	276	189	20	0.4943	0.4029	0.0644
Orlando-Kissimmee-Sanford, FL	250	141	22	0.168	0.6359	0.2684
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1,138	726	66	0.24	0.3331	0.0457
Phoenix-Mesa-Glendale, AZ	420	279	24	0.1987	0.4242	0.001

CBSA	Total Tracts	Eligible to Gentrify* Restricted Sample	# Gentrified	Mean $\%\Delta GC_{i,c}$**	Mean $\%\Delta AE_{i,c}$**	Correlations between Gentrification and Amenities
Pittsburgh, PA	574	368	51	0.296	0.1636	-0.0648
Portland-Vancouver-Hillsboro, OR-WA	356	215	16	0.2353	0.6339	0.056
Providence-New Bedford-Fall River, RI-MA	324	205	25	0.1789	0.2257	0.0665
Richmond, VA	214	139	9	0.2568	0.1902	-0.0727
Riverside-San Bernardino-Ontario, CA	361	262	25	0.1334	0.4052	0.1114
Rochester, NY	225	150	14	0.252	0.4336	0.0841
Sacramento--Arden-Arcade--Roseville, CA	325	232	19	0.1865	0.0952	-0.0269
San Antonio-New Braunfels, TX	420	273	24	0.1789	0.2214	0.0963
San Diego-Carlsbad-San Marcos, CA	217	164	10	0.6855	0.4716	0.035
San Francisco-Oakland-Fremont, CA	527	346	26	0.2438	0.3286	0.0972

CBSA	Total Tracts	Eligible to Gentrify* Restricted Sample	# Gentrified	Mean $\%\Delta GC_{i,c}$**	Mean $\%\Delta AE_{i,c}$**	Correlations between Gentrification and Amenities
San Jose-Sunnyvale-Santa Clara, CA	709	442	36	0.2805	0.337	-0.0545
Seattle-Tacoma-Bellevue, WA	294	179	20	0.3288	0.2214	-0.0065
St. Louis, MO-IL	566	343	31	0.2347	0.7702	0.0415
Tampa-St. Petersburg-Clearwater, FL	377	239	40	0.2251	0.3381	0.021
Virginia Beach-Norfolk-Newport News, VA-NC	275	192	29	0.1814	0.5077	0.1135
Washington-Arlington-Alexandria, DC-VA-MD-WV	644	444	44	0.4923	0.4746	0.0103

*** Denotes summary statistic came from the restrict sample that includes only tracts in the bottom 60% of the CBSA distribution of share of college graduates ** Denotes the summary statistics for tracts that gentrified in the restricted sample including only tracts in the bottom 60% of the CBSA distribution of share of college graduates**

Table XI: Restricted Sample-Binary vs. Continuous Gentrification Measure

	Binary Gentrification Top 40% in 2014		Binary Gentrification Top 20% in 2014		Continuous Measure	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Gentrification	ΔAE	Gentrification	ΔAE	Gentrification	ΔAE
$\Delta AE_{c,i}$	-0.0240		-0.00868		6.87e-05	
	[0.0190]		[0.00616]		[0.00837]	
$SCG_{c,i,2000}$	1.028***		0.124***		0.146***	
	[0.0219]		[0.00713]		[0.00963]	
$\%\Delta Renters_{c,i,(2007,2014)}$	-0.00139		0.00170		0.0104***	
	[0.00348]		[0.00112]		[0.00148]	
$\%\Delta UnempRate_{c,i,(2007,20}$	-0.00490**	-0.000654	-0.00101	-1.92e-05	-0.00443***	0.00574
	[0.00221]	[0.00522]	[0.000720]	[0.00533]	[0.000978]	[0.00558]
Gentrification Measure		0.181***		1.592***		1.759***
		[0.0487]		[0.413]		[0.318]
$SAE_{c,i,2000}$		-0.953***		-0.959***		-0.933***
		[0.0422]		[0.0429]		[0.0442]
Median Firm Age in 2014		-0.00151***		-0.00154***		-0.00145***
		[0.000311]		[0.000313]		[0.000310]
Constant	-0.102***	0.686***	-0.0189***	0.701***	-0.0101	0.657***
	[0.0162]	[0.0373]	[0.00526]	[0.0376]	[0.00714]	[0.0395]
Observations	16,111	16,111	16,111	16,111	16,111	16,111
Adjusted R-squared	0.134	0.073	0.025	0.034	0.055	-0.009

Note: All results from the restricted sample of only tracts whose share of college educated was in the bottom 60th percentile in 2007. Gentrification Measure is either 1 if the neighborhood moved to the top 40th percentile or the top 20th percentile in 2014 in the Binary measures. In columns (5) and (6), gentrification is simply $\%\Delta CG$. $\%\Delta AE$ is as detailed in equation (4) *** p<0.01, ** p<0.05, * p<0.1

Table XII: Unrestricted Sample: Binary vs. Continuous Gentrification Measures

	Binary Top 40%		Binary Top 20%		Continuous	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Gentrification	ΔAE	Gentrification	ΔAE	Gentrification	ΔAE
$\Delta AE_{c,i}$	-0.00841		-0.0233		-0.0205	
	[0.00835]		[0.0193]		[0.0248]	
$SCG_{c,i,2000}$	-0.0428***		0.168***		0.0325***	
	[0.00533]		[0.0123]		[0.0126]	
$\%\Delta Renters_{c,i,(2007,2014)}$	0.00714***		0.0132***		0.00887***	
	[0.00103]		[0.00231]		[0.00250]	
$\%\Delta UnempRate_{c,i,(2007,2014}$	-0.00650***	-0.000119	-0.0105***	0.00340	-0.0168***	0.0365**
	[0.000925]	[0.00454]	[0.00214]	[0.00406]	[0.00277]	[0.0143]
Gentrification Measure		0.231		0.564***		2.325***
		[0.385]		[0.119]		[0.747]
$SAE_{c,i,2000}$		-1.060***		-1.025***		-1.062***
		[0.0346]		[0.0357]		[0.0653]
Median Firm Age in 2014		-0.00142***		-0.00130***		-0.00132***
		[0.000262]		[0.000264]		[0.000289]
Constant	0.0165**	0.722***	0.132***	0.609***	0.394***	-0.209
	[0.00773]	[0.0305]	[0.0179]	[0.0395]	[0.0231]	[0.304]
Observations	26,139	26,139	26,139	26,139	26,139	26,139
Adjusted R-squared	0.024	0.067	0.011	-0.020	0.003	-2.651

Note: All results from the unrestricted sample. Gentrification Measure is either 1 if the neighborhood moved to the top 40th percentile or the top 20th percentile in 2014 in the Binary measures. In columns (5) and (6), gentrification is simply $\%\Delta CG$. $\%\Delta AE$ is as detailed in equation (4) *** p<0.01, ** p<0.05, * p<0.1

Table XIII: Falsification Tests

	Always Top 40% in CBSA ΔCG		Always Bottom 40% in CBSA ΔCG	
	Binary		Binary	
VARIABLES	(1) Gentrification	(2) ΔAE	(3) Gentrification	(4) ΔAE
$\Delta AE_{c,i}$	0.0417*		-0.0120	
	[0.0246]		[0.0307]	
$SCG_{c,i,2000}$	0.982***		-1.989***	
	[0.0205]		[0.0455]	
$\%\Delta Renters_{c,i,(2007,2014)}$	-7.81e-05		-0.00554	
	[0.00239]		[0.00743]	
$\%\Delta UnempRate_{c,i,(2007,2014)}$	-0.0103***	-0.00378	0.0103***	0.00427
	[0.00259]	[0.00549]	[0.00368]	[0.00689]
Gentrification Measure		-0.0631		-0.121***
		[0.0448]		[0.0415]
$SAE_{c,i,2000}$		-1.151***		-0.916***
		[0.0598]		[0.0524]
Median Firm Age in 2014		-0.000577		-0.00125***
		[0.000478]		[0.000370]
Constant	0.299***	0.771***	1.146***	0.760***
	[0.0279]	[0.0672]	[0.0249]	[0.0555]
Observations	10,028	10,028	10,972	10,972
Adjusted R-squared	0.194	0.062	0.166	0.077

*** p<0.01, ** p<0.05, * p<0.1

Notes: The gentrified dependent variable is measured as 0/1 and the amenities dependent variable is measure as %Δ AE as detailed in equation (4) .* p<0.01, ** p<0.05, * p<0.1. Standard errors are denoted in brackets. In columns (1) and (2), only neighborhoods whose share of college graduates was in the top 40th percentile in 2007 are considered. In columns (3) and (4), only neighborhoods whose share of college graduates was in the bottom 40th percentile in 2007 are considered.**

Chapter III: Revisiting the Burden of the Gas Tax in an Electric Vehicle World

I. Introduction

In January 2019, the United States had the fifth-largest per capita electric vehicle fleet in the world, behind Norway, Netherlands, Sweden, and Belgium.[23] In the last decade and a half, the electrification of private transportation has grown steadily. While the overall percentage of electric vehicles sold each year in the private transportation sector is still rather small, constituting about 1.15 percent of cars on the road, there have been large year-to-year increases over the past decade.[24] As depicted in Figure I, the sales of Plug-in Electric Vehicles (PEV) have quadrupled since The Transportation Research Center at Argonne National Laboratory started counting in 2011. Part of the increase in uptake of electric vehicles may be due to the national government and state governments' tax credits and subsidies for the purchase of a new electric vehicle. For example, the first 200,000 buyers of any PEV model are eligible for a federal tax credit of up to $7,500 (IRS, 2020). Assuming, each purchase of a PEV displaces a gas-powered vehicle, this growth in PEV purchases necessarily, all else equal, means less consumption of gasoline. The primary purpose of this paper is to determine whether this changing consumption pattern has resulted in a change in the elasticity of demand for gasoline and its implications for tax incidence of the national gasoline tax.

The national gas tax for unleaded gasoline is 18.4 cents per gallon sold and has remained constant for 27 years. (Shaper 2018). In 2018, approximately 22 billion dollars was collected in federal highway tax revenue (Highway Statistics 2018, Federal Highway Administration). As

---

[23] Green Car Reports, (Everts, 2019); For more information please: https://www.greencarreports.com/news/1121186_us-has-worlds-second-highest-electric-car-population

[24] CityLab, (Bellan, 2018); The Grim State of Electric Vehicle Adoption in the U.S. :https://www.citylab.com/transportation/2018/10/where-americas-charge-towards-electric-vehicles-stands-today/572857/

this tax is a flat percentage, the tax is naturally regressive at face value (i.e., poorer households who spend the same amount of money on gasoline as a richer household pay a larger share of their income in gasoline taxes). Since PEVs are typically purchased by wealthier households, the burden of the gasoline tax will increasingly fall on poorer households as PEV sales increase.[25]

To estimate the elasticity of demand for gasoline, this paper will estimate a consumer demand system taking into consideration private transportation fuel sources, namely the consumption of gasoline. To estimate the incidence of increases in the gasoline tax, I estimate two scenarios where electric vehicles enter a consumer's expenditure set. The empirical structure of this paper is based on West and Williams (2004), updating their analysis from 20 years ago in a private transportation environment that has undergone dramatic change since their work.

Given this change in purchasing habits of a common commodity, one might consider how the entrance of electric vehicles in the private transportation market has affected household's demand elasticity for gasoline and whether the gas tax is becoming increasingly regressive. In other words, as more households invest in and purchase electric vehicles, the burden of the gas tax (a primary funding source for infrastructure) falls onto the people that purchase gasoline. Furthermore, electric vehicles are often purchased by wealthier households, which leaves the burden of the national gas tax more so on lower-income households. Therefore, these wealthy households will still be using the infrastructure but no longer paying to support its upkeep. From

[25] Chakraborty et al. (2019) report that 88% of the owners of electric vehicles in California survey had incomes higher than the median income for the state. This is unsurprising as PEVs often are at least 90% more expensive in purchase price than comparable internal combustive engines. KellyBlueBook (2018). See : https://mediaroom.kbb.com/2018-02-01-Average-New-Car-Prices-Rise-Nearly-4-Percent-For-January-2018-On-Shifting-Sales-Mix-According-To-Kelley-Blue-Book

a social welfare perspective, this leads one to question the tax incidence of gasoline when fewer households consume gasoline. To my knowledge, no paper has examined how changes to a current policy like the national gas tax affects consumers in an environment where private transportation looks drastically different than it did 25 years ago.

This paper continues as follows: Section II describes this paper's place in the relevant literature; Section III describes the data; Section IV explains the methodology; Section V details the simulations and discuss; Section VI concludes.

II. Literature Review

The gasoline tax is a policy instrument that both supports local and national infrastructure as well as internalizes some negative externalities from gasoline consumption. Because taxes change the price of gasoline directly and are therefore a market based approach that let consumers change their optimal behavior, they are generally preferred to other policy instruments such as fuel standards (Davis and Knittel 2016). Also, from a welfare perspective, gas taxes are often preferred to fuel standards (Anderson et al. 2011) as they give consumers more flexibility in their choice set and typically exceed estimated per-gallon climate damages.

Since gasoline taxes are levied as a flat rate based on the purchase, they are considered a regressive tax. In other words, for the same type of car and same miles driven, a poor household would pay more in taxes as a share of their household income than a rich household. Empirically, however, the regressivity of gas taxes depends on how the revenue is used (West and Williams 2004; Bento et al. 2009). If the revenue collected is simply discarded, then the tax is regressive. However, if the revenue is recycled via a cut in the labor tax or returned via a lump sum, then the regressivity of the tax is less severe and may also be slightly progressive (West and Williams 2004).

Many analyses with a PEV focus tend to analyze how environmentally "friendly" (or unfriendly) electric cars are compared to traditional vehicles. Holland et al. (2016) find that overall, tax credit incentives towards the purchase of a PEV are not justified, as PEV are often just as "dirty" as gasoline power vehicles. Tscharaktschiew (2015), is the only paper, to my knowledge, that considers gasoline taxes in the presence of an emerging electric vehicle market. He estimates the efficient gasoline tax in a world with electric vehicles for Germany and finds that optimal gasoline taxation may increase or decrease depending on PEV diffusion, electricity generation source, and external cost of PEVs. The aforementioned paper does not address the question of what the optimal gasoline tax should be, but rather, focuses on the implications of changing demand for the regressivity of the tax and the ability of the tax to generate revenue in an environment of fewer gasoline powered vehicles. This would be the closest comparison to the work I do in this paper. However, I focus not on changing optimal taxation but seeing how the optimal gas tax from previous works affects the regressivity of the national gas tax when more households are consuming electric vehicles (i.e., decreasing their consumption of gasoline).

Finally, the expectation that the gasoline tax has become more regressive as sales of PEVs have increased assumes that it is wealthier households making most of those purchases. This assumption is supported by work by Chakraborty et al. (2019), who report that 88% of the owners of electric vehicles in their California survey had incomes higher than the median income for the state. Tal and Nicholas (2016) explore the impact of the federal tax credit towards electric vehicles across several states. Most buyers were found to have household incomes of $50,000 or higher. Figure 2 details the income distribution from their sample. As discussed in the next section, I do not have information on which households purchase electric vehicles. Therefore, I rely on the assumption that, overwhelmingly, wealthier households are the most likely owners of PEVs.

III. Data

In order to estimate a demand system, I need information on household expenditures on gasoline, wages, and other goods. In order to simulate tax increases and estimate tax incidence, I need information on state level gasoline taxes. Therefore, data for this paper comes from several publicly available sources and one private source.

The Consumer Expenditure Survey (CEX) is a nationally representative survey that contains detailed questions on household spending habits, income, hours worked, demographic, and geographic information for all individuals in the household. Households are surveyed between 1 to 4 times. This survey is used to estimate a detailed consumer demand system. Following West and Williams (2004), the sample is restricted to two types of households: households with 1 adult and households with 2 adults and their dependent or under 18 year old children. I utilize the 2016 through 2018 quarterly interview files. These interview files contain household information of expenditures such as gasoline and individual information on wages. I focus on this time period as opposed to earlier years, as the first plug-in hybrid became available in 2010 and the costs of electric battery vehicles only started to significantly decrease in 2013.[26] In other words, at least prior to 2013, I wouldn't expect any major differences in the private transportation market that would influence household's elasticity of gasoline. As seen in Figure I, the choices of PEV models has expanded since 2010, giving households several options of PEV by 2016.

[26] Department of Energey (2020); Timeline: History of the Electric Car: https://www.energy.gov/timeline/timeline-history-electric-car

The second vital aspect of a consumer demand system is information on prices. I utilize the Council for Community and Economic Research Historical Cost of Living Index, which I purchased. This data set contains quarterly price information for a bundle of goods going back to the first quarter over the time span of my CEX data. Other data used for the analyses include the quarterly state unemployment rate and state gasoline tax rates on gasoline, which are obtained from the Bureau of Labor Statistics and the Federation of Tax Administrators, respectively.[27] Unemployment rate data are important for modeling labor force selection in order to accurately estimate wages. Tax information by each state over time is needed in order to correctly estimate the change in fuel price. State taxes do not vary much over 2016-2018, with only 15 states introducing changes to their state tax rate.

Finally, information on the purchases of electric vehicles is not contained in the CEX, nor does it contain information on what kind of car a person owns. The overarching assumption I am making is that most households are consuming gasoline powered cars. I simulate electric vehicle purchasing decisions by modifying household purchases of gasoline. My analyses can't make any comments on purchases of hybrid vehicles (vehicles that consume both gasoline and electricity).

My analyses feature 1- and 2-Adult households. I further restrict my sample to include only adults between 18-64 who are working; working individuals are most likely to regularly make use of some mode of transportation. The results presented will, therefore, only be generalizable to working individuals in this age range. Table XIV presents the summary characteristics of households. There are 4,342 1-adult household level observations and 3,962 2-

[27]Bureau of Labor Statistics (2020); https://www.bls.gov/lau/rdscnp16.htm and Federation of Tax Administrators (2020); https://www.taxadmin.org/current-tax-rates

adult household observations. There are several differences between these 2016-2018 samples and the samples in West and Williams (2004) who use data between 1996-1998. My sample is younger, with an average age of 40 for both household types compared to 41 and 44 in West and Williams, even when keeping the age range the same. Respondents consumed slightly more gasoline -- 13.78 and 24.67 then vs 15.27 and 26.71 now. The sample is much more diverse with much less white non-Hispanic representation in the 2016-2018 sample.

IV. Methodology

There are several steps in order to get from the demand system to tax incidence and then simulations of electric car purchases. First, I estimate a baseline demand system. Then, I calculate measures of tax incidence, under differing distributive assumptions, when the national gas tax is increased. Finally, I then simulate PEV purchases by decreasing households purchases of gasoline. The goal is to see how the tax incidence changes as PEV purchases increase, primarily among the wealthy.

IV.A. Demand System

Because I need to incorporate behavioral responses into my measures of tax incidence, I estimate an "Almost Ideal Demand System" (AIDS) popularized Deaton and Muellbauer (1980). Like West and Williams (2004), I estimate a linear AIDS. This type of demand system relies on the expenditure function. It is detailed below:

$$\log c(\boldsymbol{p}, u) = (1-u)\log\big(\mathrm{a}(\mathbf{p})\big) + u * \log\big(\mathrm{b}(\mathbf{p})\big) \tag{1}$$

$$\log a(\boldsymbol{p}) = \alpha_0 + \sum_k \alpha_k \log(p_k) + \frac{1}{2}\sum_k \sum_j \gamma_{kj} \log p_k \log p_j \tag{2}$$

$$\log b(\boldsymbol{p}) = \log a(\boldsymbol{p}) + \beta_0 \prod\nolimits_k p_k^{\beta^k} \tag{3}$$

Equation (1) is a consumer's (household's) expenditure function. $\boldsymbol{p}$ indicates a vector of prices and u is the household's level of utility. Equations (2) and (3) are the specific functional forms used by Deaton and Muellbauer (1980). I obtain demand equations for the three goods considered -- gasoline, leisure, and all other expenditures -- by applying Shepherds Lemma. These are the resulting estimating equations.

$$s_{ih} = \alpha_i + \sum_j \gamma_{ij} \log(p_{jh}) + \beta_i \log\left(\frac{y_h}{P_h}\right) \tag{4}$$

Equation (4) is the share of the household h's income on spent on good i, which is a function of prices for all good consumed, p_j, and real income, $\frac{y_h}{P_h}$. P_h is a non-linear price index described by household h.

$$\log P_h = \alpha_0 + \sum_k \alpha_k \log(p_{kh}) + \frac{1}{2}\sum_j \sum_k \gamma_{kj} \log(p_{kh}) \log(p_{jh}) \tag{5}$$

The system can be estimated using techniques developed by Poi (2012) for the statistical software *Stata* called *quaids*. This command allows users to estimate equation (4) with demographic and other controls. Since all expenditure shares must sum to one, the following requirements must be met: α_i (the budget share's intercept) sum to 1, all γ_{ij} sum to zero, all β_i sum to zero, and $\gamma_{ij} = \gamma_{ji}$ (Slutsky Symmetry). Equation (4) is modified to include an additional vector of covariates (X_h), including standard demographic information, educational outcomes, number of children, and state fixed effects; and the family's a propensity to consume gasoline (calculated by estimating a choice model of gasoline):

$$s_{ih} = \alpha_i + \sum_j \gamma_{ij} \log(p_{jh}) + \beta_i \log\left(\frac{y_h}{P_h}\right) + \eta_i X_h + \delta_i \hat{\lambda}_h \tag{4'}$$

The propensity to consume gasoline ($\hat{\lambda}_h$) is a regressor that is constructed by estimating a probit model, using the CEX, where the observed dependent variable is equal to one if the household is observed to be consuming gasoline and zero otherwise. This allows the parameter estimates to be generalized to the population even though equation (4') is estimated only using households who consume a positive amount of gasoline West and Williams (2004).The propensity for a household to consume gasoline are not directly available in the CEX.[28] Therefore, before I estimate the demand system described by equation (4') which includes a household's propensity to consume gasoline estimated with a probit model. Net wages are calculated by using household's marginal tax rate (estimate by using NBER's Taxsim software). More details on these two preliminary steps are detailed Appendix C Tables C-I-C-III.

IV.B. Tax Incidence

Following West and Williams (2004), I estimate the model using three types of goods: gasoline, leisure, and other goods. I do not directly estimate demand for electricity. When I simulate household's switching to electric vehicles, I let the value of their expenditure share of gasoline transfer to the "other good" which includes electricity. This is because the gasoline to electricity tradeoff from switching is not known with certainty. Furthermore, Equation (4') is estimated separately for both single adult households and two-adult households (ones with a female and male adult). Households are then split up by income level, which is calculated as total expenditures. Doing so allows me to see how elasticity for gasoline and other goods changes with income and by household type.

[28] I estimate net wages to not only match West and Williams (2004) methodology but to also estimate a labor tax cut in future versions of this paper.

Next, I formulate measures of tax incidence. There are several incidence measures that can measure the tax burden. I focus on a consumer surplus with heterogeneous demand elasticities and an incidence measure where demand is unresponsive to price changes. These are two incidence measures detailed in West and Williams (2004) and presented here with slight modifications. The change in consumer surplus ($\Delta\, CS_h$) measure with heterogeneous demand elasticities takes the following form:

$$\Delta\, CS_h = \left\{ \frac{\bar{x}_g \bar{p}_h^g}{\varepsilon_h^g + 1} \left[1 - \left(\frac{p_h^g}{\bar{p}_h^g} \right)^{\varepsilon_h^g + 1} \right] \right\} + T_g \tag{6}$$

where, h is the representative household for a given quartile, $\bar{x}_g$ is mean expenditure share of gasoline before the price change, $\bar{p}_h^g$ is the mean price of gasoline before the price change, p_h^g is the mean price of gasoline after the price , ε_h^g is the uncompensated own price elasticity of demand for gasoline, and T_g is a lump sum transfer to each household from the increased revenue from increasing the gasoline tax. All I need for the above equation: (1) spending on the taxed good before the tax imposition; (2) Percent change in the price induced by the tax; (3) the lump sum transfer to each household; (4) Own price elasticity of a good whose price changes. This measure requires a demand system to be estimated.

The second incidence measure, where demand is unresponsive to price takes the following form (all terms are defined above):

$$Incidence = \left(\bar{p}_h^g - p_h^g\right)\bar{x}_h^g + T_g \tag{7}$$

This measure does not require the estimation of a demand system. This is meant to act as a baseline to the demand response/consumer surplus measure. The no-demand response is the price change less the lump sum tax (if that is the revenue policy).

V. Simulations and Discussions

V.A. Baseline Estimate

I first estimate a demand system on expenditures related to gasoline, other goods, and leisure in order to get estimates of demand elasticities to use in the simulations. For each household type, all adult members are observed to be working. The demand system estimates from equation (4') are detailed in Appendix C Tables C-II-C-III. Table XV and XVI list the uncompensated and compensated demand elasticities and cross-price elasticities for 1- and 2-Adult Households evaluated at the means of all variables used in the AIDS model for each model specification (i.e., 1-adult households in the 1st income quartile of 1-adult households)[29]. The model is estimated separately by income quartile and family type. Generally, the elasticity of gasoline is more inelastic at higher income levels. This pattern is consistent with estimates by West and Williams (2004) who find that upper-income levels have lower responses to changes in gasoline prices. However, the aforementioned work found a more consistent downward trend as household income increased, whereas, there are some instances where the middle household income quartiles have larger absolute elasticities. This is in direct contrast to Spiller et al., (2017), who find higher income households are relatively more price sensitive to changes in gasoline prices. Part of the difference between Spiller et al. (2017) and West and Williams (2004) and therefore, the current work, is that the elasticities in this paper and West and Williams (2004) are short run elasticities that do not allow for households to respond to increases in gasoline prices by switching to more fuel efficient vehicles. Contrastingly, the Spiller et al. (2017) model of calculating elasticity accounts for optimal vehicles miles travelled.

[29] Since these are not direct parameter estimates, bootstrapping would be required to obtain confidence intervals for these estimates

Once the demand system is estimated, I estimate tax incidence using equations (6) and (7). In the consumer surplus measure, the only price change I consider is the gas tax and I utilize the uncompensated demand elasticity for gasoline. I simulate an increase the national gasoline tax to \$1.39 which was the value chosen in West and Williams (2004), which is the optimal tax amount given damages from gasoline powered vehicles. This \$1.39 increase in gasoline from the current \$0.184 represents a 600% increase in the tax rate. State gasoline taxes are assumed to remain unchanged.

I follow two assumptions about what is done with gas tax revenue once it's been collected: it is either discarded or it is given back as a lump sum to all households. In the tax incidence measure where no lump sum is given, this simply means that $T_g = 0$. The baseline tax incidences are detailed in Tables XVII and XVIII. These numbers represent the incidence as a percent share of total income. These numbers are calculated using a "representative household" from each quartile, where the representative household in each quartile faces the mean prices and mean expenditures of that quartile. Consistent with West and Williams (2004), for both incidence measures, demand response and no demand response, the gasoline tax is highly regressive

As seen in Tables XVII and XVIII, welfare measures are represented as a percentage of total expenditures. The larger and more negative welfare measure, the more the representative household "loses" from the policy change. However, once revenue is recycled via a lump sum and transferred, lower quartiles actually benefit from the policy change, while households in higher quartiles still experience a welfare loss, albeit less than when revenue is disregarded.

V.B. Simulation 1: Gasoline Purchases Cut by 99% by Top Quartile

Before I present the results of the simulations, it is important to note that since I do not explicitly see electric vehicle purchases and therefore do not account for the potential include of subsidies included for purchases of electric vehicles. All simulations abstract away from any tax credit resulting from the purchase of a PEV. In this demand system, a decreased demand for gasoline increases demand for electricity. Since I am unable to calculate how dollars of gasoline translate into dollars of electricity power for a car that is driven the same amount, I simulate electric car purchases by letting gasoline consumption transfer over into the "share of other goods" purchased, which includes electricity purchases. My first simulation of electric car purchases is an extreme case where each household in the 4th quartile decreases their purchase of gasoline by 99%.[30] Therefore, this decrease simulates an almost total switch from gasoline to other goods (i.e., electricity).

Tables XIX and XX detail the tax incidence from this simulation for the 1- and 2-Adult households. While not shown here, each simulation requires a re-estimation of the demand system. Note that since consumption patterns were assumed to not change in the first three quartiles, the tax incidence when revenue is disregarded is unchanged from the baseline. However, since households in the fourth quartile are consuming virtually no gasoline, their incidence when revenue is disregarded essentially goes to zero.

Additionally, when there is no demand response, the lump sum transfer retains its progressive nature even with most of the top quartile no longer consuming gasoline but still receiving the lump sum transfer. However, when the top quartile essentially no longer consumes any gasoline but still receives a lump sum, the tax continues to be regressive. In other words,

[30] I don't take gasoline consumption to zero, since in a demand system with zero purchase of a good, observations will drop out of the estimation.

previously when revenue is recycled in the consumer surplus model, this makes a regressive tax progressive. However, if a large enough portion of the consumer base no longer consumes gasoline but is given the lump sum tax, the regressive tax remains regressive.

V.C. Simulation 2: Gasoline Purchases Cut for All Quartile

The final simulation is to decreases gasoline purchases for households in all quartiles. The bottom quartiles has 2% of its sample decrease their gasoline consumption by 99%, the second to bottom quartiles has 4% of its sample decrease their gasoline consumption by 99%, the second to the highest quartile has 6% of its sample decrease gasoline purchase by 99% and the top quartile has 8% of its sample decrease their gasoline consumption by 99%. Which households decrease gasoline consumption is determined randomly and the demand estimation and incidence estimation processes are repeated twenty-five times. The averaged results are found in Tables XXI and XXII.

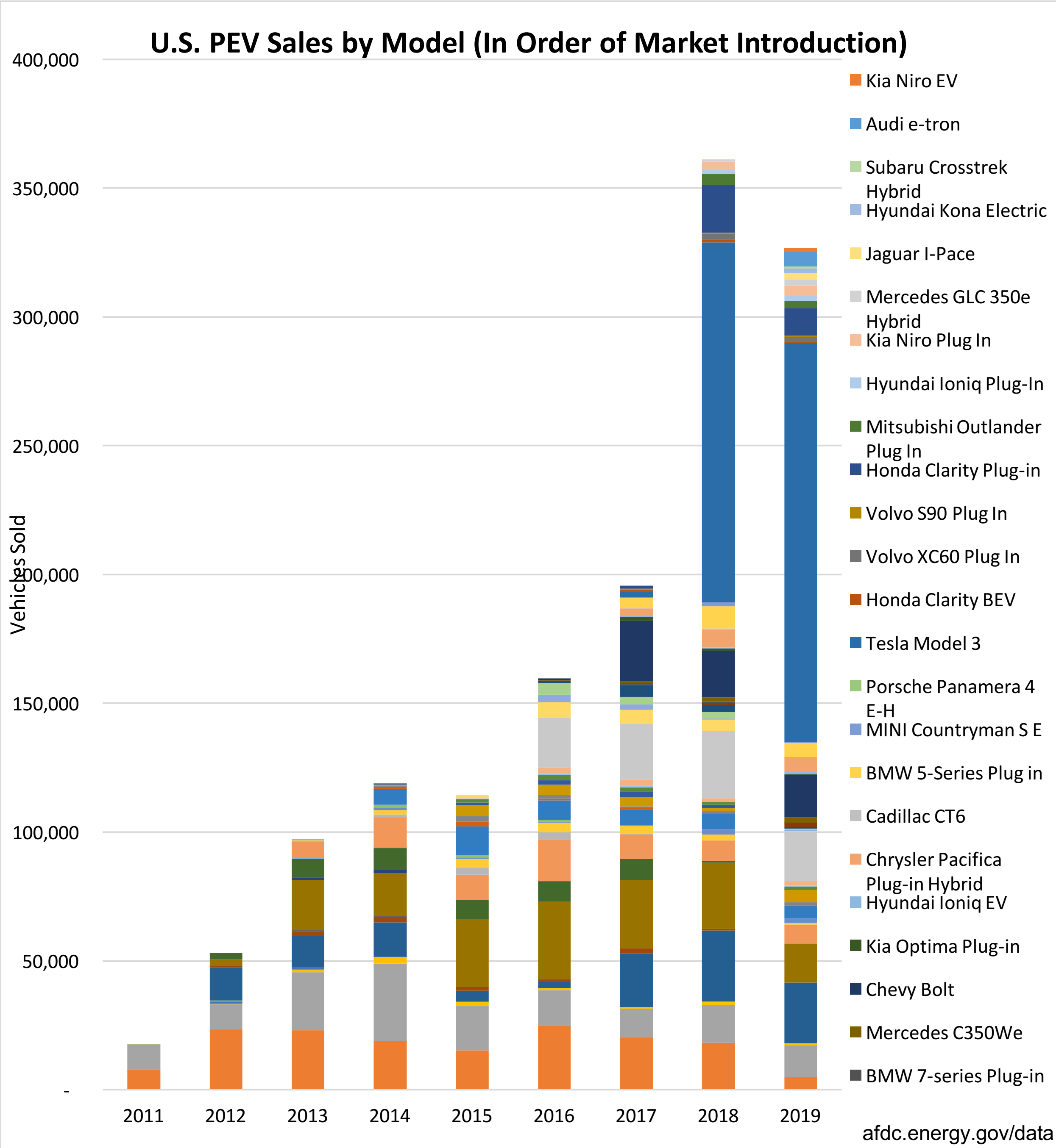
VI. Conclusion

As the nature of private transportation changes, some thought needs to be given to how existing policies, such as gas taxes, affect the consumers who still pay those taxes. As more car manufactures develop Plug-In Electric vehicles (PEVs), there will be an incentive for consumers to switch to electric vehicles. Furthermore, as climate change becomes more of a pressing concern, PEVs are often promoted as eco-friendly, therefore increasing their appeal. In this paper, I estimate a consumer demand model to illicit elasticities to calculate tax incidence from a change in the national tax from \$0.184 to \$1.39 per gallon of gasoline. I then simulate two scenarios reflective of consumers changing their purchases of gasoline-powered vehicles to electric powered ones.

The incidence of the increase in gas tax is generally regressive because it is primarily wealthy households who purchase PEVs and, hence, avoid paying the higher gasoline tax. However, this tax can become more progressive if revenue is recycled as modeled here by a lump sum transfer to households. This is important because the gas tax is currently the main source of funding for federal highways. Going forward, if the policymakers attempt to recoup lost revenue from the gas tax due to households switching over to electric vehicles, special attention needs to be paid to persons who are still paying for that tax and the burden it places on them.

Figures

Figure I: Sales of Plug-In Electric Vehicles



Source: Alternative Fuels Data Center.

Figure II: Income Distribution by PEV Purchase

Tal and Nicholas

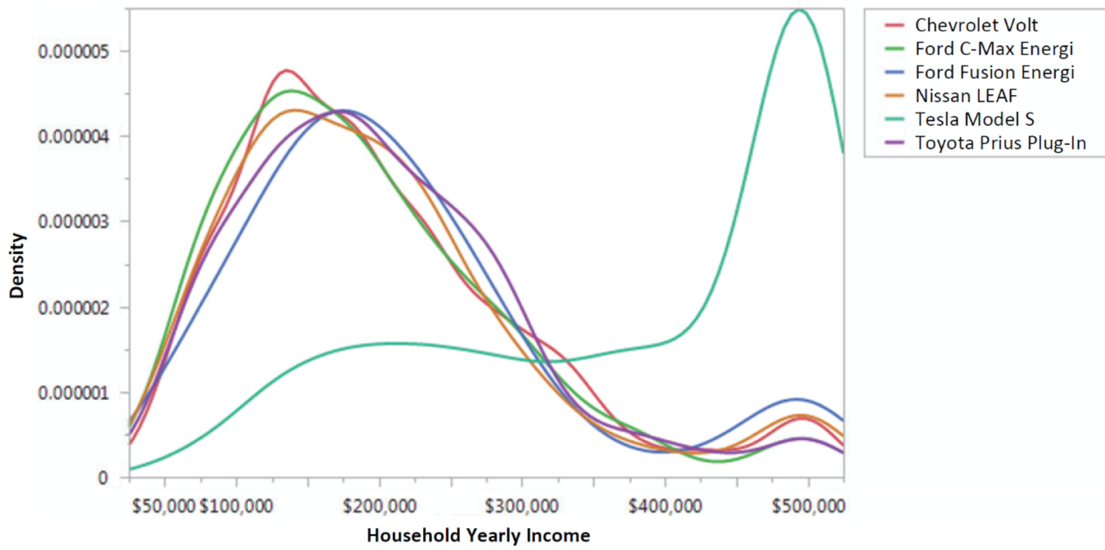


FIGURE 1 Income distribution.

Source: Tal and Nicholas (2016)

Tables

Table XIV: Chapter III Summary Statistics

Variable	1-Adult HH	2-Adult HH
Age (Age Female)	41.58	39.78
	[13.3]	[11.28]
Age Male	--	40.99
		[11.54]
Share Female	0.54	--
	[0.5]	
Log(Real Income)	6.62	7.22
	[0.51]	[0.46]
Hourly Wage (Hourly Wage Female)	10.59	8.1
	[2.94]	[2.13]
Hourly Wage Male	--	10.46
		[2.6]
Gasoline Price	2.37	2.38
	[0.41]	[0.41]
Other Good Price	1.06	1.06
	[0.19]	[0.19]
Gasoline Per Week	16.12	26.75
	[16.71]	[22.82]
Hours worked (Hours worked Female)	40.26	37.69
	[11.99]	[11.26]
Hours worked Male	--	43.84
	--	[10.52]
Share of Expenditures in Leisure (Share of Expenditures on Leisure Female)	0.59	0.28
	[0.17]	[0.09]
Share of Expenditures in Leisure Male		0.31
		[0.1]
Share Spent on Gasoline	0.02	0.02
	[0.02]	[0.01]
Share of other goods	0.39	0.4
	[0.17]	[0.15]
Number of Children	0.34	1.01
	[0.8]	[1.16]
White, NH	65.56%	69.11%

(White, NH Female)		
White, NH Male	--	70.08%
Black, NH (Black, NH Female)	15.61%	6.60%
Black, NH Male	--	7.81%
Hispanic (Hispanic Female)	6.36%	8.62%
Hispanic Male	--	7.22%
Other, NH (Other, NH Female)	12.46%	15.67%
Other Male	--	14.90%
Share less than High School (Female)	5.02%	5.02%
Share less High School Male	--	6.89%
Share High School	18.62%	17.87%
Share High School Male	--	22.19%
Share >High School (Female)	76.36%	77.12%
Share >High School Male	--	70.72%
Observations (Households)	4,983	4,047

Table XV: Baseline Demand Elasticities for 1-Adult Households

	Compensated			Uncompensated		
	Gas Price	Wage	Other Good Price	Gasoline	Wage	Other Good Price
Q=1						
Gasoline	-1.19	-0.14	1.33	-1.19	-0.08	1.35
Leisure	0.00	-0.01	0.01	-0.02	-0.99	-0.27
Other Good	0.09	0.03	-0.12	0.09	-0.04	-0.14
Q=2						
Gasoline	-0.63	-0.09	0.72	-0.63	-0.17	0.67
Leisure	0.00	-0.01	0.01	-0.03	-0.99	-0.49
Other Good	0.04	0.02	-0.06	0.04	0.00	-0.07
Q=3						
Gasoline	-0.65	-0.05	0.7	-0.66	-0.11	0.65
Leisure	0.00	-0.03	0.03	-0.03	-1.00	-0.73
Other Good	0.03	0.04	-0.07	0.03	0.00	-0.09
Q=4						

	Compensated			Uncompensated		
	Gas Price	Wage	Other Good Price	Gasoline	Wage	Other Good Price
Gasoline	0.11	-0.1	-0.01	0.12	-0.09	0.01
Leisure	-0.01	-0.23	0.23	-0.04	-0.96	-0.88
Other Good	0.00	0.15	-0.15	-0.01	-0.02	-0.42

Note: Elasticities where estimated using *quaids* command in Stata 16. Full system coefficients can be found in the appendix.
These are elasticities evaluated at the means of the variables used in the demand system estimation and therefore do not have standard errors.

Table XVI: Baseline Demand Elasticities for 2-Adult Households

	Compensated				Uncompensated			
	Gas Price	Wage Male	Wage Female	Other Good Price	Gas Price	Wage Male	Wage Female	Other Good Price
Q1								
Gasoline	-0.82	0.1	0.16	0.56	-0.82	-0.03	0.05	0.49
Male Leisure	0.00	-0.5	0.42	0.08	-0.01	-0.98	-0.01	-0.2
Female Leisure	0.01	0.46	-0.57	0.10	-0.01	-0.02	-1.01	-0.19
Other Good	0.04	0.14	0.15	-0.32	0.03	0.00	0.02	-0.41
Q2								
Gasoline	-0.18	0	-0.03	0.21	-0.19	-0.03	-0.06	0.18
Male Leisure	0.00	-0.48	0.45	0.03	-0.03	-1.00	0.00	-0.47

Female Leisure	0.00	0.52	-0.56	0.04	-0.03	0.01	-1.00	-0.45
Other Good	0.01	0.03	0.04	-0.08	0.01	-0.01	0.01	-0.11
Q3								
Gasoline	-0.91	0.02	0.04	0.86	-0.92	-0.05	-0.02	0.77
Male Leisure	0.00	-0.49	0.43	0.06	-0.03	-0.99	0.00	-0.66
Female Leisure	0.00	0.50	-0.57	0.06	-0.03	0.01	-0.99	-0.65
Other Good	0.04	0.04	0.04	-0.12	0.04	-0.01	0.00	-0.19
Q4								
Gasoline	-0.86	-0.01	-0.03	0.9	-0.87	-0.06	-0.07	0.77
Male Leisure	0	-0.64	0.27	0.37	-0.02	-0.94	0.01	-0.46

Female Leisure	0	0.31	-0.73	0.42	-0.03	0.01	-0.99	-0.41
Other Good	0.02	0.14	0.13	-0.29	0.01	-0.02	0	-0.72

Note: Elasticities where estimated using *quaids* command in Stata 16. Full system coefficients can be found in the appendix. These are elasticities evaluated at the means of the variables used in the demand system estimation and therefore do not have standard errors.

Table XVII: Baseline Estimated Tax Incidence: 1 Adult Households

Tax Incidence Measure	**Q=1**	**Q=2**	**Q=3**	**Q=4**
Revenue is Disregarded				
No Demand Response	-2.61%	-2.01%	-1.69%	-1.06%
Consumer Surplus	-4.62%	-4.30%	-3.62%	-3.37%
Revenue is Return via Lump Sum Transfer				
No Demand Response	3.94%	1.65%	0.65%	-0.03%
Consumer Surplus	1.93%	-0.64%	-1.28%	-2.33%

Note: Tax Incidence as a percentage of total expenditures

Table XVIII: Baseline Estimated Tax Incidence: 2 Adult Households

Tax Incidence Measure	**Q=1**	**Q=2**	**Q=3**	**Q=4**
Revenue is Disregarded				
No Demand Response	-2.40%	-1.89%	-1.70%	-0.94%
Consumer Surplus	-4.89%	-5.32%	-3.41%	-1.93%
Revenue is Return via Lump Sum Transfer				
No Demand Response	3.35%	1.47%	0.51%	0.07%
Consumer Surplus	0.55%	-2.14%	-1.32%	-0.98%

Note: Tax Incidence as a percentage of total expenditures

Table XIX: Gasoline Cut Top Quartile- Estimated Tax Incidence: 1 Adult Households

Tax Incidence Measure	**Q=1**	**Q=2**	**Q=3**	**Q=4**
Revenue is Disregarded				
No Demand Response	-2.61%	-2.01%	-1.69%	-0.01%
Consumer Surplus	-4.62%	-4.30%	-3.62%	-0.03%
Revenue is Return via Lump Sum Transfer				
No Demand Response	7.74%	3.77%	2.01%	1.63%
Consumer Surplus	-0.55%	-2.03%	-2.16%	0.61%

Note: Tax Incidence as a percentage of total expenditures

Table XX: Gasoline Cut Top Quartile- Estimated Tax Incidence: 2 Adult Households

Tax Incidence Measure	**Q=1**	**Q=2**	**Q=3**	**Q=4**
Revenue is Disregarded				
No Demand Response	-2.40%	-1.89%	-1.70%	-0.01%
Consumer Surplus	-4.95	-5.4	-3.42	-0.02
Revenue is Return via Lump Sum Transfer				
No Demand Response	1.36%	0.31%	-0.26%	0.65%
Consumer Surplus	-1.2	-3.2	-1.98	0.64

Note: Tax Incidence as a percentage of total expenditures

Table XXI: Gasoline Cut All Quartiles- Estimated Tax Incidence: 1 Adult Households

Tax Incidence Measure	**Q=1**	**Q=2**	**Q=3**	**Q=4**
Revenue is Disregarded				
No Demand Response	-2.55%	-1.93%	-1.59%	-0.98%
Consumer Surplus	-4.12%	-6.17%	-5.86%	-12.69%
Revenue is Return via Lump Sum Transfer				
No Demand Response	3.65%	1.53%	0.62%	0.00%
Consumer Surplus	2.05%	-2.72%	-3.65%	-12.71%

Note: Tax Incidence as a percentage of total expenditures

Table XXII: Gasoline Cut All Quartiles- Estimated Tax Incidence: 2 Adult Households

Tax Incidence Measure	**Q=1**	**Q=2**	**Q=3**	**Q=4**
Revenue is Disregarded				
No Demand Response	-2.35%	-1.82%	-1.60%	-0.87%
Consumer Surplus	-5.64%	-16.04%	-3.61%	-2.65%
Revenue is Return via Lump Sum Transfer				
No Demand Response	3.08%	1.35%	0.48%	0.09%
Consumer Surplus	-0.2%	-12.86%	-1.52%	-1.69%

Note: Tax Incidence as a percentage of total expenditures

Appendix A-Chapter I

Table A-I: Parameter estimates from the social capital equation

VARIABLES	Trust	Weak Network	Strong Network
Individual Controls			
Age	0.0453	-0.1108*	-0.2767***
	[0.0692]	[0.0641]	[0.0668]
Age squared	-0.0005	0.0030**	0.0049***
	[0.0015]	[0.0014]	[0.0015]
Age cubed	0.0324	-0.2313**	-0.2824***
	[0.1056]	[0.0983]	[0.1024]
Married=0,1	0.1594***	0.0634	-0.0693
	[0.058]	[0.0535]	[0.0549]
Number of children in HH	-0.0359	0.1275***	-0.0482*
	[0.0257]	[0.0208]	[0.0281]
Household total income GE $30,000=0,1	0.3276***	0.3806***	0.2673***
	[0.0671]	[0.0607]	[0.0598]
High school education=0,1	0.2507**	0.3495***	0.2033**
	[0.1197]	[0.1022]	[0.099]
Some college education=0,1	0.4651***	0.8726***	0.4135***
	[0.1207]	[0.1035]	[0.0971]
College graduate=0,1	0.8543***	1.422***	0.6385***
	[0.1601]	[0.1411]	[0.1449]
Hispanic=0,1	-0.6352***	-0.2978**	-0.6948***
	[0.1312]	[0.1224]	[0.1224]
Black, non-Hispanic=0,1	-0.9815***	0.1314	-0.8214***
	[0.1276]	[0.1113]	[0.1154]
Other race, non-Hispanic=0,1	-0.1867	-0.4259**	-0.7741***
	[0.198]	[0.1744]	[0.2260]
Unemployed=0,1	-0.2487**	-0.3218***	-0.2694**
	[0.1107]	[0.1122]	[0.1248]
Not in the labor force=0,1	0.0584	0.0287	0.1533**
	[0.0714]	[0.064]	[0.0618]
Citizen=0,1	0.2224*	0.6006***	0.3660***
	[0.1256]	[0.1296]	[0.1247]
Lived in area 5 yrs or less=0,1	-0.1400***	-0.2226***	-0.3043***
	[0.0505]	[0.0502]	[0.0499]
Own home=0,1	0.4438***	0.3723***	0.2480**
	[0.1145]	[0.1085]	[0.1023]

VARIABLES	Trust	Weak Network	Strong Network
Female=0,1	0.1822***	0.0079	0.0663
	[0.0498]	[0.0455]	[0.0458]
College grad * white non-Hispanic	0.3860***	0.0652	0.0194
	[0.1253]	[0.1127]	[0.1206]
College grad * own home	-0.3210***	-0.081	-0.3039***
	[0.1019]	[0.0972]	[0.0949]
White non-Hispanic * own home	-0.172	-0.1186	-0.3269***
	[0.1293]	[0.1187]	[0.1176]
Age GE 75 years	-0.0312	0.3442	-0.0288
	[0.3273]	[0.2929]	[0.2863]
Age LT 25 years	0.1904	0.3443**	0.0175
	[0.1463]	[0.1402]	[0.1521]
Regional Controls			
Live in MSA=0,1	-0.1212	-0.0131	-0.0975
	[0.0791]	[0.0727]	[0.0718]
Mid-Atlantic region=0,1	0.1511	-0.2647*	-0.0564
	[0.1676]	[0.1514]	[0.1564]
East North Central region=0,1	0.3668**	0.0125	0.0329
	[0.1829]	[0.1692]	[0.1747]
West North Central region=0,1	0.0573	-0.4085**	-0.0046
	[0.2136]	[0.1992]	[0.1960]
South Atlantic region=0,1	0.1644	0.0407	-0.1436
	[0.2215]	[0.2038]	[0.1927]
East South Central region=0,1	0.2169	-0.0174	0.0329
	[0.2444]	[0.2353]	[0.2181]
West South Central region=0,1	0.3064	0.1301	0.3535
	[0.2705]	[0.252]	[0.232]
Mountain region=0,1	0.4071*	0.0057	0.2602
	[0.2418]	[0.2272]	[0.2014]
Pacific region=0,1	0.3086	-0.0263	0.3196
	[0.2536]	[0.235]	[0.2195]
Distance Weighted Census Tract Characteristics (Exogenous Regressors)			
Share of workers in broad social capital occupations	3.127	6.101	4.541
	[5.555]	[5.176]	[4.937]
Share of workers in social	-2.111	32.38**	16.32

VARIABLES	Trust	Weak Network	Strong Network
capital industries			
	[14.85]	[14.43]	[13.71]
Labor force participation rate	2.072	-5.242	-14.81
	[9.642]	[8.767]	[9.195]
Unemployment rate	-11.72*	4.522	-11.31*
	[6.413]	[5.917]	[6.214]
Percent lived in area at least 5 years	1.664	-4.145***	-0.9244
	[1.657]	[1.429]	[1.515]
Median age	0.0995*	0.0256	-0.0509
	[0.0536]	[0.0546]	[0.0544]
Diversity index	-1.08	0.9485	-0.3008
	[0.9963]	[0.9398]	[0.9243]
Female labor force participation rate	-1.316	6.956	13.79*
	[8.155]	[7.528]	[8.028]
Percent college graduates, 25 and older	2.403	0.2479	-3.795*
	[2.536]	[2.180]	[2.275]
Population density (1000/sq mi)	-0.005	0.0398	0.0259
	[0.0436]	[0.0419]	[0.04]
Percent married households	2.162	-0.4471	0.058
	[1.924]	[1.761]	[1.723]
Percent of families with own children	6.819**	0.2388	-2.473
	[3.236]	[3.024]	[2.909]
Percent who own home	0.3233	2.308	0.7157
	[1.566]	[1.439]	[1.450]
Median household income ($00000)	-0.6543	-0.2507	1.876
	[1.722]	[1.479]	[1.476]
Observations	20,220	20,220	20,220
Pseudo R2	0.0999	0.0679	0.051

Notes: Standard errors in brackets, etc. Parameter estimates from ordered logit estimations.

Appendix B-Chapter II

Table B-I: Amenity Establishments

NAICS Description	Code	NAICS Description	Code
Supermarkets and Other Grocery (except Convenience) Stores	445110	Florists	453110
Baked Goods Stores	445291	Office Supplies and Stationery Stores	453210
Confectionery and Nut Stores	445292	Gift, Novelty, and Souvenir Stores	453220
All Other Specialty Food Stores	445299	Used Merchandise Stores	453310
Beer, Wine, and Liquor Stores	445310	Pet and Pet Supplies Stores	453910
Pharmacies and Drug Stores	446110	Art Dealers	453920
Cosmetics, Beauty Supplies, and Perfume Stores	446120	Theater Companies and Dinner Theaters	711110
Food (Health) Supplement Stores	446191	Dance Companies	711120
All Other Health and Personal Care Stores	446199	Musical Groups and Artists	711130
Men's Clothing Stores	448110	Other Performing Arts Companies	711190
Women's Clothing Stores	448120	Independent Artists, Writers, and Performers	711510
Children's and Infants' Clothing Stores	448130	Museums	712110
Family Clothing Stores	448140	Historical Sites	712120
Clothing Accessories Stores	448150	Zoos and Botanical Gardens	712130
Other Clothing Stores	448190	Nature Parks and Other Similar Institutions	712190
Shoe Stores	448210	Fitness and Recreational Sports Centers	713940
Jewelry Stores	448310	Bowling Centers	713950
Luggage and Leather Goods Stores	448320	Bed-and-Breakfast Inns	721191
Sporting Goods Stores	451110	Drinking Places (Alcoholic Beverages)	722410
Hobby, Toy, and Game Stores	451120	Full-Service Restaurants	722511
Sewing, Needlework, and Piece Goods Stores	451130	Limited-Service Restaurants	722513
Musical Instrument and Supplies Stores	451140	Cafeterias, Grill Buffets, and Buffets	722514
Book Stores	451211	Snack and Nonalcoholic Beverage Bars	722515
Beauty Salons	812112	Barber Shops	812112
Nail Salons	812113	Other Personal Care	812199
Pet Care	812910	Civic and Social Organizations	813410

NAICS Description	Code	NAICS Description	Code
Dry-cleaning and Laundering Services	812320	Movie Theaters	512131
Commercial Banking	522110	Credit Unions	522130

Table B-II: First Stage Results: OLS Estimation: Restricted Gentrification

	Bottom 60% Top 40%		Bottom 60% Top 20%	
	(1)	(2)	(3)	(4)
VARIABLES	Gentrified	$\%\Delta AE$	Gentrified	$\%\Delta AE$
$SCG_{c,i,2000}$	1.024***	0.175***	0.123***	0.175***
	[0.0211]	[0.0498]	[0.00687]	[0.0498]
$\%\Delta Renters_{c,i,(2007,2014)}$	-0.00297	0.0502***	0.000255	0.0502***
	[0.00331]	[0.00779]	[0.00108]	[0.00779]
$\%\Delta UnempRate_{c,i,(2007,2014)}$	-0.00484**	-0.00241	-0.000976	-0.00241
	[0.00221]	[0.00521]	[0.000720]	[0.00521]
$SAE_{c,i,2000}$	0.0221	-0.936***	0.00655	-0.936***
	[0.0180]	[0.0423]	[0.00584]	[0.0423]
Median Firm Age in 2014	4.63e-05	-0.00144***	7.05e-05	-0.00144***
	[0.000132]	[0.000310]	[4.29e-05]	[0.000310]
Constant	-0.118***	0.644***	-0.0250***	0.644***
	[0.0164]	[0.0387]	[0.00534]	[0.0387]
Observations	16,111	16,111	16,111	16,111
Adjusted R-squared	0.134	0.077	0.026	0.077

Note: Estimation is done via OLS. Standard errors are denoted in brackets. Gentrified=1 if the neighborhood meets the criterion described in equation (1) and (2). $\%\Delta AE$ denotes the change in amenities as described in equation (4). *** p<0.01, ** p<0.05, * p<0.1

Table B-III: First Stage Results: OLS Estimation: Unrestricted Gentrification

	Continuous Gentrification		Binary Gentrification	
	(1)	(2)	(3)	(4)
VARIABLES	Gentrified	$\%\Delta AE$	Gentrified	$\%\Delta AE$
$SCG_{c,i,2000}$	-0.0435***	0.0676***	0.0374**	0.0676***
	[0.00517]	[0.0209]	[0.0154]	[0.0209]
$\%\Delta Renters_{c,i,(2007,2014)}$	0.00693***	0.0206***	0.00736**	0.0206***
	[0.00100]	[0.00407]	[0.00301]	[0.00407]
$\%\Delta UnempRate_{c,i,(2007,2014)}$	-0.00648***	-0.00253	-0.0168***	-0.00253
	[0.000925]	[0.00375]	[0.00277]	[0.00375]
$SAE_{c,i,2000}$	0.00821	-1.015***	0.0210	-1.015***
	[0.00855]	[0.0346]	[0.0256]	[0.0346]
Median Firm Age in 2014	2.30e-05	-0.00126***	0.000116	-0.00126***
	[6.48e-05]	[0.000262]	[0.000194]	[0.000262]
Constant	0.0108	0.676***	0.377***	0.676***
	[0.00789]	[0.0320]	[0.0236]	[0.0320]
Observations	26,139	26,139	26,139	26,139
Adjusted R-squared	0.024	0.068	0.002	0.068

Note: Estimation is done via OLS. Standard errors are denoted in brackets. Continuous gentrification is the measure described in equation (3). Binary gentrification is an indicator =1 if the neighborhood's $\%\Delta CG$ is in the top 40% of the CBSA distribution of $\%\Delta CG$. $\%\Delta AE$ denotes the change in amenities as described in equation (4). This*** p<0.01, ** p<0.05, * p<0.1

Appendix C-Chapter III

Table C-I: Gasoline Probit for 1-Adult Households and 2-Adult Households

VARIABLES	1-Adult Households Gasoline Probit	2-Adult Households Gasoline Probit
Log (Total Expenditures)	0.505***	0.576***
	(0.0146)	(0.0129)
Age (or Age Female)	0.0176	-0.0530
	(0.0112)	(0.0374)
Age Male		0.00846
		(0.0345)
Age2 (or Age2 Female	-0.000209	0.000681
	(0.000136)	(0.000446)
Age2 Male		-0.000141
		(0.000414)
Black, NH (or Black, NH Female)	-0.248***	-0.152
	(0.0534)	(0.213)
Black, NH Male		-0.361**
		(0.176)
Hispanic (or Hispanic Female)	-0.340***	-0.0465
	(0.0755)	(0.146)
Hispanic Male		0.118
		(0.200)
Other, NH (or Other, NH Female)	0.0329	0.190
	(0.0648)	(0.183)
Other, NH Male		-0.0734
		(0.141)
High School (or High School Female)	0.161*	0.147
	(0.0853)	(0.175)
High School Male		0.353**
		(0.169)
Some College (or Some College Female)	0.304***	0.0962
	(0.0830)	(0.183)
Some College Male		0.0980
		(0.155)
College (or College Female)	0.270***	0.325**
	(0.0854)	(0.161)
College Male		0.0223
		(0.163)
Married	-0.168**	
	(0.0700)	
Number of Children	-0.00114	0.0358
	(0.0277)	(0.0397)
Homeownership	0.635***	0.497***

	(0.0527)	(0.0899)
Log(Price of Gasoline)	0.564*	0.0515
	(0.305)	(0.690)
Log(Price of Other Good)	-0.925***	0.509
	(0.200)	(1.627)
Female	-0.0426	
	(0.0425)	
Constant	-4.322***	-2.663***
	(0.315)	(0.800)
Observations	11,841	10,860

Note: This is done because some households may not consume gasoline. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table C-II: Almost Ideal Demand System for 1-Adult Households

Variables	Equation	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Age	Gas	-1.27e-05	-2.99e-05	-8.90e-07	4.59e-05
		(1.08e-05)	(2.20e-05)	(3.00e-05)	(4.45e-05)
Age	Leisure	0.000229**	0.000283	-1.28e-05	-0.000690**
		(0.000106)	(0.000176)	(0.000186)	(0.000281)
Age	Other	-0.000216**	-0.000253	1.37e-05	0.000644**
		(9.83e-05)	(0.000169)	(0.000184)	(0.000273)
White, NH	Gas	-0.000393	0.00133	-0.00149	0.00286*
		(0.000441)	(0.00103)	(0.00163)	(0.00146)
White, NH	Leisure	0.00674	-0.0141	0.00476	-0.0194
		(0.00469)	(0.0111)	(0.0128)	(0.0123)
White, NH	Other	-0.00635	0.0128	-0.00327	0.0166
		(0.00438)	(0.0106)	(0.0123)	(0.0120)
Black, NH	Gas	0.000976*	0.00178	-0.000355	0.00248
		(0.000511)	(0.00119)	(0.00191)	(0.00181)
Black, NH	Leisure	-0.0109**	-0.00637	-0.00203	0.00734
		(0.00540)	(0.0129)	(0.0131)	(0.0153)
Black, NH	Other	0.00994**	0.00458	0.00239	-0.00982
		(0.00503)	(0.0122)	(0.0128)	(0.0150)
Hispanic	Gas	0.000425	0.00259**	0.000165	0.00644***
		(0.000543)	(0.00114)	(0.00194)	(0.00220)
Hispanic	Leisure	-0.00180	-0.0152	0.0106	-0.00815
		(0.00563)	(0.0118)	(0.0132)	(0.0150)
Hispanic	Other	0.00138	0.0126	-0.0108	0.00171
		(0.00523)	(0.0112)	(0.0127)	(0.0145)
High School	Gas	-0.00143**	0.00142	-0.000807	0.00305
		(0.000631)	(0.00130)	(0.00352)	(0.00488)
High School	Leisure	0.00661	0.00694	-0.00731	-0.0586**
		(0.00461)	(0.00907)	(0.0126)	(0.0260)
High School	Other	-0.00518	-0.00835	0.00811	0.0556**
		(0.00423)	(0.00874)	(0.0130)	(0.0263)
Some College	Gas	-0.00146**	-0.000312	-0.00299	-0.00115
		(0.000568)	(0.00128)	(0.00347)	(0.00431)
Some College	Leisure	0.00686**	0.00919	0.0139	-0.0533**
		(0.00347)	(0.00935)	(0.0112)	(0.0258)
Some College	Other	-0.00540*	-0.00888	-0.0110	0.0544**
		(0.00319)	(0.00903)	(0.0118)	(0.0261)
College	Gas	-0.00171***	-0.00178	-0.00543	-0.00141
		(0.000600)	(0.00133)	(0.00350)	(0.00426)
College	Leisure	0.00793*	0.0140	0.0312***	-0.0295

Variables	Equation	Quartile 1	Quartile 2	Quartile 3	Quartile 4
		(0.00407)	(0.00976)	(0.0117)	(0.0249)
College	Other	-0.00622	-0.0123	-0.0258**	0.0309
		(0.00379)	(0.00937)	(0.0121)	(0.0253)
Female	Gas	-1.48e-05	-0.000857	-0.000879	-0.000995
		(0.000316)	(0.000587)	(0.000763)	(0.00108)
Female	Leisure	-0.00346	-0.00146	-0.0253***	-0.0153**
		(0.00331)	(0.00476)	(0.00468)	(0.00637)
Female	Other	0.00348	0.00231	0.0262***	0.0163***
		(0.00307)	(0.00449)	(0.00461)	(0.00618)
Real Income	Gas	0.00131***	0.00327**	0.00810***	0.000897
		(0.000335)	(0.00166)	(0.00219)	(0.00145)
Real Income	Leisure	-0.0409***	-0.143***	-0.261***	-0.198***
		(0.00455)	(0.00826)	(0.00920)	(0.0193)
Real Income	Other	0.0396***	0.140***	0.253***	0.198***
		(0.00440)	(0.00810)	(0.00915)	(0.0195)
Number of Children	Gas	0.000469***	0.000600	-0.000258	-0.000298
		(0.000142)	(0.000473)	(0.000838)	(0.000870)
Number of Children	Leisure	-0.00356***	0.00106	0.0134***	0.0160***
		(0.00133)	(0.00327)	(0.00394)	(0.00574)
Number of Children	Other	0.00309**	-0.00166	-0.0132***	-0.0157***
		(0.00124)	(0.00314)	(0.00382)	(0.00563)
Gas Propensity	Gas	-0.00225**	-0.00131	0.00372	-0.00123
		(0.00111)	(0.00342)	(0.00522)	(0.0107)
Gas Propensity	Leisure	0.0170**	-0.0523**	-0.0730**	-0.248***
		(0.00851)	(0.0254)	(0.0311)	(0.0760)
Gas Propensity	Other	-0.0147*	0.0536**	0.0693**	0.249***
		(0.00787)	(0.0243)	(0.0307)	(0.0734)
α	Gas	0.0831***	0.0496***	0.0420***	0.0314***
		(0.0110)	(0.00832)	(0.00664)	(0.0105)
α	Leisure	-0.0948**	-0.00572	-0.00511	-0.0398***
		(0.0380)	(0.00606)	(0.00321)	(0.00962)
α	Other	1.012***	0.956***	0.963***	1.008***
		(0.0355)	(0.0103)	(0.00777)	(0.0129)
β	Gas	-0.0218***	-0.0343***	-0.0656***	-0.0318*
		(0.00352)	(0.0127)	(0.0163)	(0.0187)
β	Leisure	0.455***	1.278***	2.229***	1.950***
		(0.0374)	(0.0555)	(0.0653)	(0.173)
β	Other	-0.433***	-1.243***	-2.163***	-1.918***

Variables	Equation	Quartile 1	Quartile 2	Quartile 3	Quartile 4
		(0.0364)	(0.0542)	(0.0652)	(0.177)
γ	Gas, Gas	-0.00409	0.00568	0.00581	0.0204
		(0.00539)	(0.00660)	(0.00695)	(0.0128)
γ	Gas, Leisure	0.00106	-0.00328***	-0.00220**	0.000448
		(0.000810)	(0.00111)	(0.000914)	(0.00113)
γ	Gas, Other	0.00302	-0.00240	-0.00361	-0.0209
		(0.00544)	(0.00678)	(0.00695)	(0.0129)
γ	Leisure, Leisure	-0.0182**	0.0109	0.00639	-0.0224***
		(0.00715)	(0.00713)	(0.00478)	(0.00816)
γ	Leisure Other	0.0171**	-0.00760	-0.00418	0.0219***
		(0.00669)	(0.00663)	(0.00462)	(0.00779)
γ	Other, Other	-0.0202**	0.01000	0.00779	-0.00103
		-1.27e-05	(0.00920)	(0.00864)	(0.0147)
Observations		1,252	1,243	1,247	1,241

Notes: Includes state and year fixed effects not shown here. Robust standard errors are clustered by households. *** p<0.01, ** p<0.05, * p<0.1

Table C-III: Almost Ideal Demand System for 2-Adult Households

Variables	Good	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Age Female	Gas	0.000344**	-0.000294*	-0.000580**	0.000198**
		(0.000142)	(0.000171)	(0.000228)	(8.17e-05)
Age Female	Leisure-Male	-0.00385***	-0.00265**	0.00272**	-0.000666
		(0.000979)	(0.00109)	(0.00108)	(0.000414)
Age Female	Leisure-Female	0.00322***	0.00309***	0.00315***	-2.75e-05
		(0.000928)	(0.000993)	(0.00100)	(0.000369)
Age Female	Other	0.000284	-0.000152	-0.00529***	0.000496
		(0.000864)	(0.00128)	(0.00114)	(0.000639)
Age Male	Gas	-0.000780***	1.47e-05	0.000193	4.02e-05
		(0.000157)	(0.000169)	(0.000224)	(7.98e-05)
Age Male	Leisure-Male	0.00922***	0.00529***	0.00560***	-8.03e-05
		(0.000935)	(0.00106)	(0.00106)	(0.000403)
Age Male	Leisure-Female	0.00138	-0.00335***	-0.000797	-0.000486
		(0.000878)	(0.000981)	(0.000985)	(0.000357)
Age Male	Other	-0.00982***	-0.00196	-0.00500***	0.000526
		(0.000724)	(0.00122)	(0.00109)	(0.000620)
Age2 Female	Gas	-6.21e-06***	2.84e-06	6.86e-06**	-2.07e-06**
		(1.82e-06)	(2.16e-06)	(2.74e-06)	(9.40e-07)
Age2 Female	Leisure-Male	5.86e-05***	3.42e-05**	-2.61e-05**	7.75e-06
		(1.21e-05)	(1.39e-05)	(1.31e-05)	(4.77e-06)
Age2 Female	Leisure-Female	-2.47e-05**	-3.54e-05***	-3.56e-05***	-3.20e-07
		(1.14e-05)	(1.26e-05)	(1.21e-05)	(4.25e-06)
Age2 Female	Other	-2.77e-05**	-1.66e-06	5.49e-05***	-5.36e-06
		(1.08e-05)	(1.63e-05)	(1.40e-05)	(7.37e-06)
Age2 Male	Gas	1.13e-05***	3.53e-07	-2.55e-06	-6.05e-07
		(2.12e-06)	(2.07e-06)	(2.65e-06)	(9.06e-07)
Age2 Male	Leisure-Male	-0.000124***	-6.30e-05***	-7.04e-05***	1.47e-06
		(1.18e-05)	(1.31e-05)	(1.27e-05)	(4.54e-06)
Age2 Male	Leisure-Female	-3.42e-05***	3.85e-05***	9.65e-06	6.67e-06*
		(1.10e-05)	(1.20e-05)	(1.17e-05)	(4.01e-06)
Age2 Male	Other	0.000147***	2.41e-05	6.33e-05***	-7.53e-06
		(9.25e-06)	(1.51e-05)	(1.31e-05)	(6.96e-06)

Variables	Good	Quartile 1	Quartile 2	Quartile 3	Quartile 4
White, NH Female	Gas	-0.00283***	-0.00191**	-0.00266	-9.56e-05
		(0.000645)	(0.000949)	(0.00168)	(0.000550)
White, NH Female	Leisure-Male	0.00886**	-0.0107**	0.00894	8.65e-05
		(0.00414)	(0.00493)	(0.00680)	(0.00193)
White, NH Female	Leisure-Female	0.0217***	0.00782	-0.00571	0.00164
		(0.00386)	(0.00476)	(0.00657)	(0.00178)
White, NH Female	Other	-0.0278***	0.00475***	-0.000579	-0.00163
		(0.00184)	(0.00163)	(0.00286)	(0.00184)
White, NH Male	Gas	0.00262***	-0.000847	0.000542	0.000585
		(0.000629)	(0.000835)	(0.00131)	(0.000420)
White, NH Male	Leisure-Male	-0.00862**	-0.00265	0.00254	-0.000673
		(0.00413)	(0.00433)	(0.00529)	(0.00147)
White, NH Male	Leisure-Female	-0.0121***	0.00351	-0.00142	-0.000155
		(0.00386)	(0.00419)	(0.00511)	(0.00136)
White, NH Male	Other	0.0181***	-1.24e-05	-0.00166	0.000243
		(0.00179)	(0.00142)	(0.00223)	(0.00141)
Black, NH Female	Gas	-0.00259***	-0.00119	-0.00348**	-0.000307
		(0.000605)	(0.000933)	(0.00169)	(0.000551)
Black, NH Female	Leisure-Male	-0.0143***	-0.0144***	0.00266	-0.00166
		(0.00440)	(0.00484)	(0.00680)	(0.00192)
Black, NH Female	Leisure-Female	0.0195***	0.0143***	0.00209	0.00325*
		(0.00398)	(0.00468)	(0.00657)	(0.00178)
Black, NH Female	Other	-0.00264	0.00127	-0.00127	-0.00128
		(0.00302)	(0.00159)	(0.00286)	(0.00183)
Black, NH Male	Gas	0.000940	-0.00158*	-0.000661	0.000104
		(0.000630)	(0.000839)	(0.00134)	(0.000417)
Black, NH Male	Leisure-Male	0.00777*	-0.00340	0.0139**	0.00352**
		(0.00436)	(0.00433)	(0.00549)	(0.00146)
Black, NH Male	Leisure-Female	-0.0103**	-0.0114***	-0.00991*	-0.000361

Variables	Good	Quartile 1	Quartile 2	Quartile 3	Quartile 4
		(0.00407)	(0.00416)	(0.00528)	(0.00135)
Black, NH Male	Other	0.00156	0.0164***	-0.00328	-0.00326**
		(0.00270)	(0.00222)	(0.00309)	(0.00140)
Hispanic Female	Gas	-0.00254***	-0.00189**	-0.00435**	-0.000141
		(0.000659)	(0.000959)	(0.00169)	(0.000555)
Hispanic Female	Leisure-Male	-0.0371***	-0.0227***	-0.00323	-0.00280
		(0.00470)	(0.00502)	(0.00687)	(0.00203)
Hispanic Female	Leisure-Female	0.0312***	0.0261***	0.0196***	0.00435**
		(0.00429)	(0.00484)	(0.00664)	(0.00186)
Hispanic Female	Other	0.00843**	-0.00154	-0.0120***	-0.00141
		(0.00340)	(0.00233)	(0.00345)	(0.00220)
Hispanic Male	Gas	-4.92e-05	-0.00266***	-0.00160	-0.000130
		(0.000687)	(0.000864)	(0.00137)	(0.000418)
Hispanic Male	Leisure-Male	0.0273***	0.0129***	0.0210***	0.00524***
		(0.00481)	(0.00436)	(0.00541)	(0.00157)
Hispanic Male	Leisure-Female	-0.0187***	-0.0183***	-0.0164***	-0.00128
		(0.00447)	(0.00422)	(0.00524)	(0.00144)
Hispanic Male	Other	-0.00852***	0.00800***	-0.00292	-0.00383**
		(0.00277)	(0.00144)	(0.00230)	(0.00174)
High School Female	Gas	0.00172*	-0.00153	-0.00117	0.000157
		(0.00100)	(0.000942)	(0.00117)	(0.000252)
High School Female	Leisure-Male	-0.0413***	-0.000692	0.0103**	-0.00384***
		(0.00647)	(0.00488)	(0.00481)	(0.00106)
High School Female	Leisure-Female	0.0126**	0.0168***	0.0137***	0.00352***
		(0.00611)	(0.00472)	(0.00463)	(0.000934)
High School Female	Other	0.0271***	-0.0146***	-0.0228***	0.000165
		(0.00315)	(0.00160)	(0.00279)	(0.00136)
High School Male	Gas	-0.000859	0.00101	-0.000317	0.000373
		(0.00102)	(0.00103)	(0.00129)	(0.000258)

Variables	Good	Quartile 1	Quartile 2	Quartile 3	Quartile 4
High School Male	Leisure-Male	0.0300***	0.0171***	0.00542	0.00270***
		(0.00711)	(0.00541)	(0.00520)	(0.000879)
High School Male	Leisure-Female	-0.0230***	-0.00450	-0.00937*	-0.00280***
		(0.00668)	(0.00521)	(0.00503)	(0.000805)
High School Male	Other	-0.00617*	-0.0136***	0.00427*	-0.000272
		(0.00359)	(0.00247)	(0.00220)	(0.000839)
Some College Female	Gas	0.00425***	0.00151	0.00200	0.000205
		(0.00134)	(0.00138)	(0.00174)	(0.000526)
Some College Female	Leisure-Male	-0.0246***	0.0138*	0.00949	-0.00328*
		(0.00860)	(0.00718)	(0.00700)	(0.00185)
Some College Female	Leisure-Female	-0.0276***	-0.0116*	-0.00296	0.000206
		(0.00803)	(0.00694)	(0.00676)	(0.00171)
Some College Female	Other	0.0480***	-0.00361	-0.00853***	0.00287
		(0.00386)	(0.00236)	(0.00295)	(0.00177)
Some College Male	Gas	0.000618	-0.000435	0.000352	0.000899*
		(0.00131)	(0.00138)	(0.00170)	(0.000525)
Some College Male	Leisure-Male	-0.0181**	-0.0178**	-0.0107	0.00160
		(0.00905)	(0.00716)	(0.00686)	(0.00184)
Some College Male	Leisure-Female	-0.00810	0.0143**	-0.000184	-0.000768
		(0.00847)	(0.00693)	(0.00661)	(0.00170)
Some College Male	Other	0.0255***	0.00394*	0.0105***	-0.00173
		(0.00399)	(0.00235)	(0.00289)	(0.00177)
College Female	Gas	0.000201	-0.00115	0.00158	-0.000158
		(0.000994)	(0.00103)	(0.00138)	(0.000328)

Variables	Good	Quartile 1	Quartile 2	Quartile 3	Quartile 4
College Female	Leisure-Male	-0.0177***	-0.0117**	0.00350	-0.00299***
		(0.00685)	(0.00536)	(0.00554)	(0.00115)
College Female	Leisure-Female	0.0145**	0.0167***	0.00563	0.00279***
		(0.00652)	(0.00517)	(0.00536)	(0.00106)
College Female	Other	0.00304	-0.00382**	-0.0107***	0.000363
		(0.00376)	(0.00177)	(0.00236)	(0.00110)
College Male	Gas	0.00131	0.00258**	0.000661	0.000493
		(0.00108)	(0.00113)	(0.00154)	(0.000359)
College Male	Leisure-Male	-0.00102	0.0121**	0.000326	0.00165
		(0.00750)	(0.00586)	(0.00623)	(0.00126)
College Male	Leisure-Female	-0.0168**	-0.0120**	-0.00924	-0.00277**
		(0.00705)	(0.00567)	(0.00602)	(0.00116)
College Male	Other	0.0165***	-0.00276	0.00825***	0.000632
		(0.00343)	(0.00192)	(0.00264)	(0.00124)
Real Income	Gas	0.00373***	0.00241**	0.00800***	-5.91e-05
		(0.000618)	(0.00117)	(0.00193)	(0.000213)
Real Income	Leisure-Male	-0.0553***	-0.0838***	-0.126***	-0.0214***
		(0.00433)	(0.00618)	(0.00756)	(0.000948)
Real Income	Leisure-Female	-0.0499***	-0.0572***	-0.116***	-0.0181***
		(0.00401)	(0.00588)	(0.00728)	(0.000823)
Real Income	Other	0.101***	0.139***	0.234***	0.0396***
		(0.00306)	(0.00363)	(0.00674)	(0.00151)
Number of Children	Gas	0.000180	0.000376	-0.000664*	-7.07e-05
		(0.000148)	(0.000238)	(0.000386)	(8.30e-05)
Number of Children	Leisure-Male	-0.00316***	0.00433***	0.00562***	0.000160
		(0.00102)	(0.00129)	(0.00153)	(0.000415)
Number of Children	Leisure-Female	0.00202**	0.00211*	0.0119***	0.00108***
		(0.000996)	(0.00122)	(0.00148)	(0.000372)
Number of Children	Other	0.000954	-0.00681***	-0.0168***	-0.00117*
		(0.000705)	(0.000944)	(0.00134)	(0.000644)

Variables	Good	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Gas Propensity	Gas	-0.00795***	-0.00415	-0.0101	0.000152
		(0.00259)	(0.00411)	(0.00839)	(0.00286)
Gas Propensity	Leisure-Male	-0.0103	-0.00921	0.0195	0.00489
		(0.0179)	(0.0213)	(0.0337)	(0.0101)
Gas Propensity	Leisure-Female	0.00749	-0.00712	-0.0195	0.00538
		(0.0168)	(0.0205)	(0.0326)	(0.00934)
Gas Propensity	Other	0.0108	0.0205***	0.0101	-0.0104
		(0.00821)	(0.00729)	(0.0145)	(0.00991)
α	Gas	0.00233	0.0129**	0.000892	0.000895
		(0.00453)	(0.00614)	(0.00738)	(0.00722)
α	Leisure-Male	-0.00300***	-0.000217	-0.000361	0.00392***
		(0.000590)	(0.000515)	(0.000538)	(0.00125)
α	Leisure-Female	-0.00131**	-0.000980**	-1.54e-05	0.00316**
		(0.000514)	(0.000475)	(0.000536)	(0.00124)
α	Other	0.00198	-0.0117*	-0.000516	-0.00797
		(0.00450)	(0.00615)	(0.00742)	(0.00734)
β	Gas	0.0287***	-0.00228	-0.00492***	-0.0244***
		(0.00311)	(0.00157)	(0.00107)	(0.00279)
β	Leisure-Male	0.0126***	0.000139	-0.00427**	-0.0302***
		(0.00242)	(0.000962)	(0.00195)	(0.00290)
β	Leisure-Female	-0.0383***	0.00236**	0.00955***	0.0507***
		(0.00196)	(0.00117)	(0.00253)	(0.00460)
β	Other	0.0140***	-0.000517	-0.00191**	-0.0254***
		(0.00245)	(0.00112)	(0.000957)	(0.00452)
γ	Gas, Gas	-0.0252***	0.00136	0.00619***	0.0525***
		(0.00166)	(0.00112)	(0.00213)	(0.00653)
γ	Gas, Leisure Male	0.0616***	0.00798	-0.0152*	-0.0952***
		(0.00489)	(0.00627)	(0.00860)	(0.0124)
γ	Gas, Leisure Female	0.0440***	0.0418***	0.0444***	0.0770***
		(0.00564)	(0.00734)	(0.00752)	(0.0165)

Variables	Good	Quartile 1	Quartile 2	Quartile 3	Quartile 4
γ	Gas Leisure Other	0.148***	-0.0108***	-0.00853***	-0.234***
		(0.0103)	(0.00277)	(0.00243)	(0.0158)
γ	Leisure Male, Leisure Male	0.130***	0.00301	-0.00376	-0.174***
		(0.00973)	(0.00292)	(0.00271)	(0.0177)
γ	Leisure Male, Leisure Female	0.678***	0.966***	0.968***	1.331***
		(0.0142)	(0.00810)	(0.00832)	(0.0260)
γ	Leisure Male, Other	-0.0294***	-0.0208**	-0.0593***	-0.0163***
		(0.00510)	(0.00972)	(0.0158)	(0.00562)
γ	Leisure Female, Leisure Female	0.405***	0.713***	0.925***	0.270***
		(0.0329)	(0.0458)	(0.0562)	(0.00824)
γ	Leisure Female, Other	0.346***	0.533***	0.956***	0.224***
		(0.0305)	(0.0438)	(0.0552)	(0.00873)
γ	Other, Other	-0.722***	-1.226***	-1.822***	-0.478***
		(0.0200)	(0.0245)	(0.0470)	(0.0101)
Observations		1,017	1,009	1,012	1,009

Bibliography

Allen, W. David. 2000. "Social Networks and Self-Employment." *The Journal of Socio-Economics* 29 (5): 487–501. https://doi.org/10.1016/S1053-5357(00)00086-X.

Anderson, Soren T., Ian W. H. Parry, James M. Sallee, and Carolyn Fischer. 2011. "Automobile Fuel Economy Standards: Impacts, Efficiency, and Alternatives." *Review of Environmental Economics and Policy* 5 (1): 89–108. https://doi.org/10.1093/reep/req021.

Azoulay, Pierre, Benjamin Jones, J. Daniel Kim, and Javier Miranda. 2018. "Age and High-Growth Entrepreneurship." w24489. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w24489.

Bauernschuster, Stefan, Oliver Falck, and Stephan Heblich. 2010. "Social Capital Access and Entrepreneurship." *Journal of Economic Behavior & Organization* 76 (3): 821–33. https://doi.org/10.1016/j.jebo.2010.09.014.

Baum-Snow, Nathaniel, and Daniel Hartley. 2020. "Accounting for Central Neighborhood Change, 1980–2010." *Journal of Urban Economics* 117 (May): 103228. https://doi.org/10.1016/j.jue.2019.103228.

Baum-Snow, Nathaniel, and Daniel A. Hartley. 2016. "Accounting for Central Neighborhood Change, 1980-2010." SSRN Scholarly Paper ID 2839686. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2839686.

Bento, Antonio M., Lawrence H. Goulder, Mark R. Jacobsen, Von Haefen, and Roger H. 2009. "Distributional and Efficiency Impacts of Increased US Gasoline Taxes." *American Economic Review* 99 (3): 667–99. https://doi.org/10.1257/aer.99.3.667.

Borjas, George J. 1986. "The Self-Employment Experience of Immigrants." *The Journal of Human Resources* 21 (4): 485–506. https://doi.org/10.2307/145764.

Brummet, Quentin, and Davin Reed. 2019. "The Effects of Gentrification on the Well-Being and Opportunity of Original Resident Adults and Children." Working paper (Federal Reserve Bank of Philadelphia) 19–30. Federal Reserve Bank of Philadelphia. https://doi.org/10.21799/frbp.wp.2019.30.

Cavaglia, Chiara. 2015. "A New Cross-Country Investigation on the Patterns of Intergenerational Mobility." SSRN Scholarly Paper ID 2624799. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2624799.

Chakraborty, Debapriya, David S. Bunch, Jae Hyun Lee, and Gil Tal. 2019. "Demand Drivers for Charging Infrastructure-Charging Behavior of Plug-in Electric Vehicle Commuters." *Transportation Research Part D: Transport and Environment* 76 (November): 255–72. https://doi.org/10.1016/j.trd.2019.09.015.

Couture, Victor, and Jessie Handbury. 2019. "Urban Revival in America, 2000 to 2010." w24084. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w24084.

Cox, Danial A., and Ryan Streeter. 2019. "The Importance of Place: Neighborhood Amenities as a Source of Social Connection and Trust." *American Enterprise Institute - AEI* (blog). May 2019. https://www.aei.org/research-products/report/the-importance-of-place-neighborhood-amenities-as-a-source-of-social-connection-and-trust/.

Currie, Janet, and Aaron Yelowitz. 2000. "Are Public Housing Projects Good for Kids?Q." *Journal of Public Economics*, 26.

Davidsson, Per, and Benson Honig. 2003. "The Role of Social and Human Capital among Nascent Entrepreneurs." *Journal of Business Venturing* 18 (3): 301–31. https://doi.org/10.1016/S0883-9026(02)00097-6.

Davis, Lucas W., and Christopher R. Knittel. 2016. "Are Fuel Economy Standards Regressive?" Working Paper 22925. National Bureau of Economic Research. https://doi.org/10.3386/w22925.

Deaton, Angus, and John Muellbauer. 1980. "An Almost Ideal Demand System." *The American Economic Review* 70 (3): 312–26.

Dee, Thomas S., and William N. Evans. 2003. "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates." *Journal of Labor Economics* 21 (1): 178. https://doi.org/10.1086/344127.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64 (5): 1001–44.

Ding, Lei, and Jackelyn Hwang. 2016. "The Consequences of Gentrification: A Focus on Residents' Financial Health in Philadelphia." *Cityscape* 18 (3): 27–56.

Ding, Lei, Jackelyn Hwang, and Eileen Divringi. 2016. "Gentrification and Residential Mobility in Philadelphia." *Regional Science and Urban Economics* 61 (November): 38–51. https://doi.org/10.1016/j.regsciurbeco.2016.09.004.

Edlund, Lena, Cecilia Machado, and Maria Micaela Sviatschi. 2019. "Gentrification and the Rising Returns to Skill." w21729. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w21729.

Faberman, R. Jason. 2011. "The Relationship Between the Establishment Age Distribution and Urban Growth*." *Journal of Regional Science* 51 (3): 450–70. https://doi.org/10.1111/j.1467-9787.2010.00703.x.

Federal Highway Administration. 2020. "Table FE-9 | Highway Statistics 2018 - Policy | Federal Highway Administration." February 19, 2020. https://www.fhwa.dot.gov/policyinformation/statistics/2018/fe9.cfm.

Findeis, Jill L., and Leif Jensen. 1998. "Employment Opportunities in Rural Areas: Implications for Poverty in a Changing Policy Environment." *American Journal of Agricultural Economics* 80 (5): 1000–1007. https://doi.org/10.2307/1244194.

Glaeser, Edward L. 2007. "Entrepreneurship and the City." Working Paper 13551. National Bureau of Economic Research. https://doi.org/10.3386/w13551.

Glaeser, Edward L., Hyunjin Kim, and Michael Luca. 2018. "Nowcasting Gentrification: Using Yelp Data to Quantify Neighborhood Change." *AEA Papers and Proceedings* 108 (May): 77–82. https://doi.org/10.1257/pandp.20181034.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78 (6): 1360–80.

Harpham, Trudy, Emma Grant, and Elizabeth Thomas. 2002. "Measuring Social Capital within Health Surveys: Key Issues." *Health Policy and Planning* 17 (1): 106–11. https://doi.org/10.1093/heapol/17.1.106.

Hartley, Daniel. 2013. "Gentrification and Financial Health." *Economic Trends*, November. https://www.clevelandfed.org/newsroom and events/publications/economic trends/2013 economic trends/et 20131106 gentrification and financial health.

Hellerstein, Judith K., David Neumark, and Melissa McInerney. 2008. "SPATIAL MISMATCH OR RACIAL MISMATCH?" *Journal of Urban Economics* 64 (2): 464–79. https://doi.org/10.1016/j.jue.2008.04.003.

Henderson, Jason. 2002. "Building the Rural Economy with High-Growth Entrepreneurs." *Economic Review*, no. Q III: 45–70.

Henig, Jeffrey R. 1980. "Gentrification and Displacement Within Cities: A Comparative Analysis." *Social Science Quarterly (University of Texas Press)* 61 (3/4): 638–52.

Holland, Stephen P., Erin T. Mansur, Nicholas Z. Muller, and Andrew J. Yates. 2016. "Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors." *American Economic Review* 106 (12): 3700–3729. https://doi.org/10.1257/aer.20150897.

Holzer, Harry J. 1991. "The Spatial Mismatch Hypothesis: What Has the Evidence Shown?" *Urban Studies* 28 (1): 105–22. https://doi.org/10.1080/00420989120080071.

Hotchkiss, Julie L., and Anil Rupasingha. 2018a. "The Role of Social Capital in Migration Decisions." Working Paper. Federal Reserve Bank of Atlanta.

———. 2018b. "Wage Determination in Social Occupations: The Role of Individual Social Capital." In *Transitions through the Labor Market*, 46:127–81. Research in Labor Economics 46. Emerald Publishing Limited. https://doi.org/10.1108/S0147-912120180000046005.

Hwang, Jackelyn, and Robert J. Sampson. 2014. "Divergent Pathways of Gentrification: Racial Inequality and the Social Order of Renewal in Chicago Neighborhoods." *American Sociological Review* 79 (4): 726–51. https://doi.org/10.1177/0003122414535774.

IRS. 2020. "Plug In Electric Vehicle Credit IRC 30 and IRC 30D | Internal Revenue Service." 2020. https://www.irs.gov/businesses/plug-in-electric-vehicle-credit-irc-30-and-irc-30d.

Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *The Quarterly Journal of Economics* 82 (2): 175–97. https://doi.org/10.2307/1885893.

Kerr, Sari Pekkala, and William R. Kerr. 2016. "Immigrant Entrepreneurship." Working Paper 22385. National Bureau of Economic Research. https://doi.org/10.3386/w22385.

Klevmarken, Anders. 1982. "Missing Variables and Two-Stage Least-Squares Estimation from More than One Data Set." Working Paper 62. IUI Working Paper. https://www.econstor.eu/handle/10419/95205.

Kolko, Jed. 2007. "The Determinants of Gentrification." SSRN Scholarly Paper ID 985714. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=985714.

Kwon, Seok-Woo, Colleen Heflin, and Martin Ruef. 2013. "Community Social Capital and Entrepreneurship." *American Sociological Review* 78 (6): 980–1008.

Lancee, Bram. 2010. "The Economic Returns of Immigrants' Bonding and Bridging Social Capital: The Case of the Netherlands." *The International Migration Review* 44 (1): 202–26.

Lannoo, Steven, Pieter-Paul Verhaeghe, Bart Vandeputte, and Carl Devos. 2012. "Differences in Social Capital Between Urban and Rural Environments." *Journal of Urban Affairs* 34 (4): 373–94. https://doi.org/10.1111/j.1467-9906.2011.00592.x.

Le, Anh T. 1999. "Empirical Studies of Self-Employment." *Journal of Economic Surveys* 13 (4): 381–416. https://doi.org/10.1111/1467-6419.00088.

Lester, T. William, and Daniel A. Hartley. 2014. "The Long Term Employment Impacts of Gentrification in the 1990s." *Regional Science and Urban Economics* 45 (March): 80–89. https://doi.org/10.1016/j.regsciurbeco.2014.01.003.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. 2019. "IPUMS National Historical Geographic Information System: Version 14.0 [Database]." IPUMS. http://doi.org/10.18128/D050.V14.0.

Martin, Isaac William, and Kevin Beck. 2018. "Gentrification, Property Tax Limitation, and Displacement." *Urban Affairs Review* 54 (1): 33–73. https://doi.org/10.1177/1078087416666959.

McKinnish, Terra, Randall Walsh, and T. Kirk White. 2010. "Who Gentrifies Low-Income Neighborhoods?" *Journal of Urban Economics* 67 (2): 180–93. https://doi.org/10.1016/j.jue.2009.08.003.

Meltzer, Rachel. 2016. "Gentrification and Small Business: Threat or Opportunity?" *Cityscape: A Journal of Policy Development and Research* 18 (3): 29.

Meltzer, Rachel, and Pooya Ghorbani. 2017. "Does Gentrification Increase Employment Opportunities in Low-Income Neighborhoods?" *Regional Science and Urban Economics* 66 (September): 52–73. https://doi.org/10.1016/j.regsciurbeco.2017.06.002.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. "Internal Migration in the United States." *Journal of Economic Perspectives* 25 (3): 173–96. https://doi.org/10.1257/jep.25.3.173.

Moretti, Enrico. 2004. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94 (3): 656–90. https://doi.org/10.1257/0002828041464623.

Nicoletti, Cheti, and John F Ermisch. 2008. "Intergenerational Earnings Mobility: Changes across Cohorts in Britain." *The B.E. Journal of Economic Analysis & Policy* 7 (2). https://doi.org/10.2202/1935-1682.1755.

Nieto, Mariano, and Nuria González-Álvarez. 2016. "Social Capital Effects on the Discovery and Exploitation of Entrepreneurial Opportunities." *International Entrepreneurship and Management Journal* 12 (2): 507–30. https://doi.org/10.1007/s11365-014-0353-0.

Patrick, Carlianne, Heather Stephens, and Amanda Weinstein. 2016. "Where Are All the Self-Employed Women? Push and Pull Factors Influencing Female Labor Market Decisions." *Small Business Economics* 46 (3): 365–90. https://doi.org/10.1007/s11187-015-9697-2.

Pink-Harper, Stephanie A. 2015. "Educational Attainment: An Examination of Its Impact on Regional Economic Growth." *Economic Development Quarterly* 29 (2): 167–79.

Poi, Brian P. 2012. “Easy Demand-System Estimation with Quaids:” *The Stata Journal*, September. https://doi.org/10.1177/1536867X1201200306.

Price, Gregory N. 2012. “RACE, TRUST IN GOVERNMENT, AND SELF-EMPLOYMENT.” *The American Economist* 57 (2): 171–87.

Putnam, Robert. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.

Ridder, Geert, and Robert Moffitt. 2007. “Chapter 75 The Econometrics of Data Combination.” In *Handbook of Econometrics*, edited by James J. Heckman and Edward E. Leamer, 6, Part B:5469–5547. Elsevier. https://doi.org/10.1016/S1573-4412(07)06075-8.

Rupasingha, Anil, Stephan J. Goetz, and David Freshwater. 2006. “The Production of Social Capital in US Counties.” *The Journal of Socio-Economics*, Essays on Behavioral Economics, 35 (1): 83–101. https://doi.org/10.1016/j.socec.2005.11.001.

Sabatini, Fabio. 2015. “Review of ‘Immigrant Networks and Social Capital.’” *Journal of Economic Literature* 53 (2): 371–74.

Shaper, David. 2018. “It’s Been 25 Years Since The Federal Gas Tax Went Up.” NPR.Org. October 5, 2018. https://www.npr.org/2018/10/05/654670146/its-been-25-years-since-the-federal-gas-tax-went-up.

Spiller, Elisheba, Heather M. Stephens, and Yong Chen. 2017. “Understanding the Heterogeneous Effects of Gasoline Taxes across Income and Location.” *Resource and Energy Economics* 50 (November): 74–90. https://doi.org/10.1016/j.reseneeco.2017.07.002.

Stam, Wouter, Souren Arzlanian, and Tom Elfring. 2014. “Social Capital of Entrepreneurs and Small Firm Performance: A Meta-Analysis of Contextual and Methodological Moderators.” *Journal of Business Venturing* 29 (1): 152–73. https://doi.org/10.1016/j.jbusvent.2013.01.002.

Stephens, Heather, and Mark Partridge. 2011. “Do Entrepreneurs Enhance Economic Growth in Lagging Regions?” *Growth and Change* 42 (November): 431–65. https://doi.org/10.1111/j.1468-2257.2011.00563.x.

Su, Yichen. 2018. “The Rising Value of Time and the Origin of Urban Gentrification.” SSRN Scholarly Paper ID 3216013. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3216013.

Tal, Gil, and Michael Nicholas. 2016. “Exploring the Impact of the Federal Tax Credit on the Plug-In Vehicle Market.” *Transportation Research Record* 2572 (1): 95–102. https://doi.org/10.3141/2572-11.

Tscharaktschiew, Stefan. 2015. “How Much Should Gasoline Be Taxed When Electric Vehicles Conquer the Market? An Analysis of the Mismatch between Efficient and Existing Gasoline Taxes under Emerging Electric Mobility.” *Transportation Research Part D: Transport and Environment* 39 (August): 89–113. https://doi.org/10.1016/j.trd.2015.06.007.

U.S. Census Bureau. n.d. “Relationship Files.” The United States Census Bureau. Accessed May 14, 2020. https://www.census.gov/geographies/reference-files/2010/geo/relationship-files.html.

Vigdor, Jacob L., Douglas S. Massey, and Alice M. Rivlin. 2002. “Does Gentrification Harm the Poor? [With Comments].” *Brookings-Wharton Papers on Urban Affairs*, 133–82.

West, Sarah E., and Roberton C. Williams. 2004. “Estimates from a Consumer Demand System: Implications for the Incidence of Environmental Taxes.” *Journal of Environmental Economics and Management*, Including Special Symposium Section from the National Bureau of Economic Research Conference on Advances in Empirical Environmental Policy Research, 47 (3): 535–58. https://doi.org/10.1016/j.jeem.2003.11.004.

Zellner, Arnold, and H. Theil. 1962. “Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations.” *Econometrica* 30 (1): 54–78. https://doi.org/10.2307/1911287.

Vita

Kalee Elise Burns graduated with her Ph.D. in Economics from the Andrew Young School of Policy Studies in 2020. She specializes in urban, labor, and public economics. While at Georgia State University she received the Atlanta Federal Reserve Bank Fellowship and the Andrew Young School Dissertation Fellowship. Prior to her time at GSU, Kalee received a BS in Mathematics and Economics from the University of West Florida in 2014. Kalee is originally from Pensacola, Florida. In August, she will join the U.S. Census Bureau’s Social, Economic, and Housing Statistics Division as an Economist.